

1	1460.2	43.4	1557	22	AA159285	Human polynucleotide
2	1316.6	39.1	1382	22	AA161071	Human polynucleotide
3	803	23.9	1258	22	AA501149	Interferon induced
4	803	23.9	1270	22	AA501151	Interferon induced
5	663	19.7	3692	22	AA501150	Interferon induced
6	663	19.7	3704	22	AA501152	Interferon induced
7	577.2	17.2	609	22	AAH3745	Human colon cancer
8	507.2	15.1	595	22	AA140897	Probe #9583 used t
9	455.2	13.5	499	22	AA54039	Probe #22725 used
10	376.4	11.0	458	21	AA64601	Nucleotide sequen
11	371.2	11.0	352	22	AA66485	Novel human polyn

12	365.8	10.9	374	20	AAVB69625
13	356.8	10.6	387	22	AAFB66499
14	285.4	8.5	301	16	AAQ90050
15	281.4	8.4	292	22	AAI1913
16	281.4	8.4	292	22	AAI45108
17	281.4	8.4	292	22	AAI05627
18	269.2	8.0	377	15	AAO76640
19	252.8	7.5	460	22	AAI10666
20	252.8	7.5	460	22	AAI31919
21	252.8	7.5	460	22	AAI00596
22	229.6	6.8	554	22	AAI41271
23	217	6.4	217	22	AAI54310
24	194.4	5.8	450	22	AAH15201
25	178	5.3	2613	22	AAH15107
26	174	5.2	386	22	AAI14562
27	167	5.0	167	22	AAI24291
28	163.4	4.9	1600	21	AACT7190
29	145.8	4.3	936	22	AAFB58252
30	145.8	4.3	936	22	AAFB58254
31	145.8	4.3	936	22	AAFB58257
32	145.8	4.3	936	22	AAFB58259
33	145.8	4.3	936	22	AAFB58262
34	145.8	4.3	938	22	AAFB58255
35	145	4.3	145	22	AAI21760
36	144.6	4.3	936	22	AAFB58252
37	144.6	4.3	936	22	AAFB58254
38	144.6	4.3	936	22	AAFB58257
39	144.6	4.3	936	22	AAFB58259
40	144.6	4.3	936	22	AAFB58262
41	144.6	4.3	938	22	AAFB58255
42	141.8	4.2	160	15	AAO76554
43	124.6	3.7	193	16	AAH090049
44	109.6	3.3	3026	22	AAH013139
45	71.6	2.1	713	22	AAH03175

XX		
PR	29-NOV-2000;	2000US-0727344.
PR	19-OCT-2000;	2000US-0693036.
PR	14-SEP-2000;	2000US-0662191.
PR	03-AUG-2000;	2000US-0653450.
PR	19-JUL-2000;	2000US-0620312.
PR	09-JUL-2000;	2000US-0598042.
PR	25-APR-2000;	2000US-0552317.
PR	21-JAN-2000;	2000US-0488725.
PF	26-DEC-2000;	2000MO-US34263.
XX		
XX		
PD	26-JUL-2001.	
PN		wO200153312-Al.
XX		
OS	Homo sapiens.	
XX		
KW	leukaemia; ss.	
KM	cyanokinctic; thrombolytic; drug screening; arthritls; inflammation;	
KM	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KM	peripheral nervous system; neuropathy; central nervous system; CNS;	
KX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
DE	Human polynucleotide SEQ ID NO 1488.	
DT	22-OCT-2001 (first entry)	
AC	AA159285;	
ID	AA159285 standard; cDNA; 1557 BP.	
RESOLUT		

(HSE-) HXSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D:
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J:
 PI Zhao Q, Zhou P, Goodrich R, Drmanac RT:
 XX WPI: 2001-442253/47.
 DR P-PSDB: AAM40129.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Claim 1: SEQ ID NO 1488; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM442213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 1557 BP; 581 A; 248 C; 320 G; 408 T; 0 other;

Query Match 43.4%; Score 1460.2; DB 22; Length 1557;

Best Local Similarity 99.4%; Pred. No. 0; Mismatches 8; Indels 1; Gaps 1;

Matches 1476; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1882 gatttttggaactcaacctatgaacatgggcatccaatggaaaaaagctggcaaa 1941
 DB 61 gatttttggaactcaacctatgaacatgggcatccaatggaaaaaagctggcaaa 120
 QY 1942 aaaggaatcgcgaagacgtgttctgcagacacattggaagaaatgacaaagcccta 2001
 DB 121 gaaagaaatcgcgaagacgtgttctgcagacacattggaagaaatgacaaagcccta 180
 QY 2002 caaatlaagacacattcgaatgataatgatactcattcgaacattctataat 2061
 DB 181 caaatlaagacacattcgaatgataatgatactcattcgaacattctataat 240
 QY 2062 gaaaggaagataaagaaatcttcacatcagaatgataatgataatgataatgataat 2121
 DB 241 gaaaggaagataaagaaatcttcacatcagaatgataatgataatgataatgataat 300
 QY 2122 ggaatctgtaatgtaatgataatgataatgataatgataatgataatgataatgataat 2181
 DB 301 ggaatctgtaatgtaatgataatgataatgataatgataatgataatgataatgataat 360
 QY 2182 acagata-gattttcgaacttattttttgaaaaacataaatgctggaagctggc 2240
 DB 361 acagatacattctcgaacttattttttgaaaaacataaatgctggaagctggc 420
 QY 2241 tgaaaaccagataatgaaaaatgaaaagctgacacaaatgaaatgataatgataatgataat 2300
 DB 421 tgaaaaccagataatgaaaaatgaaaagctgacacaaatgaaatgataatgataatgataat 480
 QY 2301 atatactgagactgagaaatcagcagcagaaataatcttcaaaaacagcagagatgc 2360
 DB 481 atatactgagactgagaaatcagcagcagaaataatcttcaaaaacagcagagatgc 540
 QY 2361 atatgcgtcttcacgtgatactgataatgataatgataatgataatgataatgataatgataat 2420
 DB 541 atatgcgtcttcacgtgatactgataatgataatgataatgataatgataatgataatgataat 600

QY 2421 ccaccatctgattggaactgagacacagcagatgattcaaacctatgacacagatgaca 2480
 DB 601 ccaccatctgattggaactgagacacagcagatgattcaaacctatgacacagatgaca 660
 QY 2481 aaaagaagctatgataatctgcactggaataaaatcaactctgtattcgtacacagt 2540
 DB 661 aaaagaagctatgataatctgcactggaataaaatcaactctgtattcgtacacagt 720
 QY 2541 ggcgaataaagctgataatgataatgataatgataatgataatgataatgataatgataatgataat 2600
 DB 721 ggcgaataaagctgataatgataatgataatgataatgataatgataatgataatgataatgataat 780
 QY 2601 caatgaaatgacatggtccagccgtggtcgaagccagatgataatgataatgataatgataatgataat 2660
 DB 781 caatgaaatgacatggtccagccgtggtcgaagccagatgataatgataatgataatgataatgataat 840
 QY 2661 cctggtctgacagatggtcgaagccgtggtcgaagccagatgataatgataatgataatgataatgataat 2720
 DB 841 cctggtctgacagatggtcgaagccgtggtcgaagccagatgataatgataatgataatgataatgataat 900
 QY 2721 gaaatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2780
 DB 901 gaaatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 960
 QY 2781 taagatttggaaatgacatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2840
 DB 961 taagatttggaaatgacatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1020
 QY 2841 tatgccaagatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2900
 DB 1021 tatgccaagatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1080
 QY 2901 tggtaacgtctgtcgggaagatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2960
 DB 1081 tggtaacgtctgtcgggaagatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1140
 QY 2961 gaccccaaatccaagaaatcttcaatgtaaaagaaacaaacatgccaagaaatg 3020
 DB 1141 gaccccaaatccaagaaatcttcaatgtaaaagaaacaaacatgccaagaaatg 1200
 QY 3021 tgcgcactatcaataaagtgataatgataatgataatgataatgataatgataatgataatgataatgataat 3080
 DB 1201 tgcgcactatcaataaagtgataatgataatgataatgataatgataatgataatgataatgataatgataat 1260
 QY 3081 gatggtgacaaagcttgaatgctgtctcaaaatgaagaaatgataatgataatgataatgataatgataatgataat 3140
 DB 1261 gatggtgacaaagcttgaatgctgtctcaaaatgaagaaatgataatgataatgataatgataatgataatgataat 1320
 QY 3141 caaaatcaatcaaaagaaacaaatcaaaagtgataatgataatgataatgataatgataatgataatgataatgataat 3200
 DB 1321 caaaatcaatcaaaagaaacaaatcaaaagtgataatgataatgataatgataatgataatgataatgataatgataat 1380
 QY 3201 caatctgactatcaaaagtgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 3260
 DB 1381 caatctgactatcaaaagtgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1440
 QY 3261 tttctttaaatactacagtttaaacatttaataatgataatgataatgataatgataatgataatgataatgataatgataat 3320
 DB 1441 tttctttaaatactacagtttaaacatttaataatgataatgataatgataatgataatgataatgataatgataatgataat 1500
 QY 3321 ctacagaactgacataaagaaatcaataatgataatgataatgataatgataatgataatgataatgataatgataat 3365
 DB 1501 ctacagaactgacataaagaaatcaataatgataatgataatgataatgataatgataatgataatgataatgataat 1545

RESULT 2
 AAI61071
 ID AAI61071 standard; cDNA; 1382 BP.
 XX
 AC AAI61071:
 XX
 DT 22-OCT-2001 (first entry)

XX		Human polynucleotide SEQ ID NO 5060.
DE		
XX		
KW		Human; nocotropic; immunosuppressant; cyostatic; gene therapy; cancer;
XX		
KW		peripheral nervous system; neuropathy; central nervous system; CNS;
XX		
KW		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
RN		amorphotic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM		chemokineic; thrombolytic; drug screening; arthritis; inflammation;
KW		leukemia; ss.
XX		
OS	Homo sapiens.	
XX		
FN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000MO-US34263.	
XX		
DR	21-JAN-2000; 2000US-0488725,	
XX		
PR	25-APR-2000; 2000US-0552317,	
XX		
FR	09-JUL-2000; 2000US-0598042,	
XX		
PR	19-JUL-2000; 2000US-0620312,	
XX		
PR	03-AUG-2000; 2000US-0653450,	
XX		
PR	14-SEP-2000; 2000US-0662191,	
XX		
PR	19-OCT-2000; 2000US-0693036,	
XX		
XX	29-NOV-2000; 2000US-0727344.	
RA	(HYSE-) HYSEQ INC.	
PI	Tang YN, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
PI	Zhao Qn, Zhou P, Goodrich R, Dimanac RT;	
XX		
DR	WPI: 2001-442253/47.	
XX		
PT	P-PADB: AAM41915.	
XX		
PT	Noxel nucleic acids and polypeptides, useful for treating disorders	
XX		
PS	such as central nervous system injuries -	
XX		
CC	Claim 1: SEQ ID NO 5060; 1007Bpp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC	the encoded polypeptides (AAM38642-AAAM42213) with nocropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropsychic and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: immune system suppression.	
CC	Activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
CC		
XX		
SQ	Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 other:	
Query Match	39.1%; Score 1316.6; DB 22; Length 1382;	
Best Local Similarity	99.1%; Pred. No. 0;	
Matches 1355; Conservative	0; Mismatches 9; Indels 3; Gaps 3	
DY	2002 caaatatagcacacaacttcgaattgatagatgctgtatacattcatcttgaaaacctttcataat 2061 	
Db	13 caaatatagcacacaacttcgaattgatagatgcgtatatccattcatcttgaaaacctttcataat 72	
OY	2062 gaagaagaagaatagaagtlltggcaatcgatcagaagaatgatagatagatagtglttgtatg 2121 	
Db	73 gaagaagaagaatagaagtlltggcaatcgatcagaagaatgatagatagatagtglttgtatg 132	

[illegible]

Db	1213	ccaattcttgatccatccgaatgctggtttatttgatgatgagattacgaacttgattgaaq	1272			
Qy	3260	attcttttaaaactatcgtcttaaacatttaataatgatatgattatgaattt-caltta	3318			
Db	1273	attctttttaaadaatactatccatgcttaaacatttaatatgattatgattatgattccataa	1332			
Qy	3319	tgctacagaactgacataagaatcaataaataatgtgttttactctg	3365			
Db	1333	tgctacagaactgacataagaatcaataaataatgtgttttactctg	1379			
RESULT 3						
AASeq1149						
ID	AASeq1149	standard; DNA; 1258 BP.				
XX	AC					
XX	AASeq1149;					
XX	DT	12-SEP-2001 (first entry)				
Interferon induced nucleic acid, IFN4.						
KM	Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;					
KM	immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;					
KM	graft rejection; viral infection; hepatitis; aplastic anemia; cancer;					
KM	human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;					
KM	hematologic disease; chronic neutropenia; myocardial infarction;					
KM	neurological disease; Alzheimer's disease; Parkinson's disease; tumour;					
XX	amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN4; ds.					
XX						
OS	Homo sapiens.					
XX						
XX	Key	Location/Qualifiers				
FT	CDS	1..915				
FT		/*tag- a				
FT		/product- "IFN4"				
FT		/note- "Interferon induced polypeptide"				
XX	MO200118208-A2.					
XX						
XX	15-MAR-2001.					
PD						
XX						
PE	08-SEP-2000; 2000MO-US24704.					
XX						
PR	08-SEP-1999; 99US-0152921.					
PR	20-OCT-1999; 99US-0160575.					
PR	20-JAN-2000; 2000US-0177104.					
PR	07-SEP-2000; 2000US-0656633.					
(CURA-) CITRAGEN CORP.						
(BIOJ) BIOGEN INC.						
XX						
XX	Peyman JA, Da Silva A, Hochman P, Hsu A;					
PI						
DR	MPI: 2001-235201/24.					
XX	P-PSDB; AAU00296.					
XX						
PT	New interferon induced polypeptides and polynucleotides, useful for the					
PT	diagnosis, prevention and treatment of immunological, cell					
PT	proliferative disorders, such as lupus erythematosus, cancer, stroke					
PT	and Alzheimer's disease -					
XX						
XX						
PS	Claim 9; Page 29-32; 134pp; English.					
CC						
CC	The sequence represents interferon induced nucleic acid, IFN4. IFN					
CC	nucleic acids and polypeptides are useful for treating or preventing a					
CC	pathology associated with IFN polypeptide in a human. They are useful for					
CC	determining the presence of or predisposition to a disease associated					
CC	with altered levels of IFN polypeptide or polynucleotide. IFN nucleic					
CC	acids, polypeptides and antibodies are useful for diagnosis, prevention					
CC	or treatment of variety of immunological and cell proliferative					
CC	disorders, such as autoimmune diseases e.g. lupus erythematosus,					
CC	immunodeficiency diseases such as acquired immunodeficiency syndrome					

CC (AIDS), graft rejection, viral infections including hepatitis and human
CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,
CC haematologic diseases such as aplastic anaemia and chronic neutropenia
CC and cancer. In addition they are also useful for treating or diagnosing
CC various disorders associated with cell death, including myocardial
CC infarction, stroke, neurological diseases including Alzheimer's and
CC Parkinson's diseases, amotrophic lateral sclerosis and spinal muscular
CC atrophy. IFN nucleic acids and polypeptides are also useful for
CC identifying interferon-like proteins and interferon agonists, for
CC screening drugs and compounds which inhibit or enhance IFN activity or
CC function and as targets for the identification of small molecules that
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal
CC cell or tumour cell growth in mammals, including humans.
XX
Sequence 1258 BP; 399 A; 267 C; 276 G; 316 T; 0 other;

Query Match	23.98;	Score 803;	DB 22;	length 1256;
Best Local Similarity	95.5%;	Pred. No. 3.8e-189;		
Matches 853; Conservative	0;	Mismatches 0;	Indels 40;	Gaps 1

[illegible]

QY 885 catgagaaatcatcatcagaatcatttgcagattctctgtagttcag 937
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 867 catgagaaatcatcatcagaatcatttgcagattctctgtagttcag 919

RESULT 5
 AAS01150
 ID AAS01150 standard; DNA: 3692 BP.
 AC AAS01150;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Interferon induced nucleic acid, IFN5.
 XX
 KM Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;
 KM immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;
 KM graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;
 KM human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;
 KM haematologic disease; chronic neutropenia; myocardial infarction;
 KM neurological disease; Alzheimer's disease; Parkinson's disease; tumour;
 KM amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN5; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..804
 FT /tag_a
 FT /product-"IFN5"
 FT /note-"Interferon induced polypeptide"
 FT
 PN W0200118208-A2.
 PD 15-MAR-2001.
 XX
 PP 08-SEP-2000; 2000MO-US24704.
 XX
 PR 08-SEP-1999; 9905-0152921.
 PR 20-OCT-1999; 9905-0160575.
 PR 20-JAN-2000; 2000US-0177104.
 PR 07-SEP-2000; 2000US-0656633.
 XX
 PA (CUBA-) CUBAGEN CORP.
 PA (BIOJ) BIOGEN INC.
 XX
 PI Peyman JA, Da Silva A, Hochman P, Hsu A;
 XX
 WP1: 2001-235201/24.
 P-PSDB: AAO00297.
 XX
 PT New interferon induced polypeptides and polynucleotides, useful for the
 PT diagnosis, prevention and treatment of immunological, cell
 PT proliferative disorders, such as lupus erythematosus, cancer, stroke
 PT and Alzheimer's disease
 PT
 PS Claim 9: Page 33-39; 134pp; English.
 PS
 CC The sequence represents interferon induced nucleic acid, IFN5. IFN
 CC nucleic acids and polypeptides are useful for treating or preventing a
 CC pathology associated with IFN polypeptide in a human. They are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic
 CC acids, polypeptides and antibodies are useful for diagnosis, prevention
 CC or treatment of variety of immunological and cell proliferative
 CC disorders, such as autoimmune diseases e.g. lupus erythematosus,
 CC immunodeficiency diseases such as acquired immunodeficiency syndrome
 CC (AIDS), graft rejection, viral infections including hepatitis and human
 CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,
 CC haematologic diseases such as aplastic anaemia and chronic neutropenia
 CC and cancer. In addition they are also useful for treating or diagnosing
 CC various disorders associated with cell death, including myocardial
 CC infarction, stroke, neurological diseases including Alzheimer's and

CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular
 CC atrophy. IFN nucleic acids and polypeptides are also useful for
 CC identifying IFN nucleic acid like proteins and interferon agonists, for
 CC screening drugs and compounds which inhibit or enhance IFN activity or
 CC function and as targets for the identification of small molecules that
 CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal
 CC cell or tumour cell growth in mammals, including humans.
 CC
 SO Sequence 3692 BP; 1166 A; 709 C; 806 G; 1011 T; 0 other;

Query Match 19.7% Score 663, DB 22: Length 3692;
 Best Local Similarity 92.9% Pred. No. 3.3e-154;
 Matches 722; Conservative 0; Mismatches 15; Indels 40; Gaps 1;

QY 45 gcacagagcgttaacacctgtctcttctaaatggcagcagcagcagcagcattca 104
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 55 gcacagagcgttaacacctgtctcttctaaatggcagcagcagcagcagcattca 114
 QY 105 cctgtccgcagacacacagccatctgtctggagagacacctctctctgaagaag 164
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 115 cctgtccgcagacacacacagccatctgtctggagagacacctctctctgaagaag 174
 QY 165 aaagatgtcgaaatgggtattccacagacagagaatttcgcatactcatctgtgtcag 224
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 175 aaagatgtcgaaatgggtattccacagacagagaatttcgcatactcatctgtgtcag 234
 QY 225 ggcagagtgaaatgtatcatccagcagcagcagcagcagcagcagcagcagcagc 284
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 235 ggcagagtgaaatgtatcatccagcagcagcagcagcagcagcagcagcagcagc 296
 QY 285 tgcagagtgaaagcagacatctcagagacagcagcagcagcagcagcagcagcagcagc 344
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 257 --cagagtgaaagcagacatctcagagacagcagcagcagcagcagcagcagcagcagc 314
 QY 345 tgaactgtctgtgacaccttggagaaagagctctgacaccttggctgtgtgtgtgtgtgt 404
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 315 tgaactgtctgtgacaccttggagaaagagctctgacaccttggctgtgtgtgtgtgtgt 374
 QY 405 cgttgagagcccttcggagaaacccgagcagcagcagcagcagcagcagcagcagcagc 464
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 375 cgttgagagcccttcggagaaacccgagcagcagcagcagcagcagcagcagcagcagc 434
 QY 465 cagcagctgcctctccatcagcttggagaaagcagcagcagcagcagcagcagcagcagcagc 524
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 435 cagcagctgcctctccatcagcttggagaaagcagcagcagcagcagcagcagcagcagcagc 494
 QY 525 cctccttcagcccaactctgtgtgacaaagctttagttagagagcgttggataagtgcat 584
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 495 cctccttcagcccaactctgtgtgacaaagctttagttagagagcgttggataagtgcat 554
 QY 585 ggaaggaagaaacttgcacattgagaaatgagaaacccgagctgtgcctgcagaaacaaatg 644
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 555 ggaaggaagaaacttgcacattgagaaatgagaaacccgagctgtgcctgcagaaacaaatg 614
 QY 645 aaatgaatcaggtgtgaaagagctactaaagagatgtgtgagaaagaaactgtgtctc 704
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 615 aaatgaatcaggtgtgaaagagctactaaagagatgtgtgagaaagaaactgtgtctc 674
 QY 705 tgcatttctgaatgtcttctgtcacaacagaaacaaatgaactgttccaaagttaagag 764
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 675 tgcatttctgaatgtcttctgtcacaacagaaacaaatgaactgttccaaagttaagag 734
 QY 765 ctctgattgtccaaagaaatgagagattgaaatllacaaatllatgtgtc 821
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 735 ctctgattgtccaaagaaatgagagattgaaatllacaaatllatgtgtc 791

RESULT 6
 AAS01152
 ID AAS01152 standard; DNA: 3704 BP.
 XX
 AC AAS01152;


```

|||||
Db 535 TCGGATATTCACAGACGAGAAATTCGGCTATCTGCTGCTTCCAGGGCCAGGCTGAA 476
Oy 237 aatgtacatccagtgagtgagctgtgctgactgaactgaccttctgctcagaagtgaa 296
Db 475 AATGTACATCCAGCTGGAGCCCTGTGCTGACTGACTGACCTTTCTGCTCGACAGAGTGA 416
Oy 297 ggaagacattcagaagacagctccacctccggagacatgacaggaattgaactgtact 356
Db 415 GGAGGAGATTTCAGAGACAGAGTCGCCACTCCGGGAAACATGACGAGCATTTGAACCTGCT 356
Oy 357 gaggcccttgaggaaggaagcttgacacttgattgactcgggaattcgtggaagccct 416
Db 355 GAGGACCTTTGAGAGAGGAGTGTGCGACCTTGCTGACTGGGAAATTCGAGAGCCCT 296
Oy 417 ccgaggaacccgagcagcccttgccgcgcgcgtacatagaacctgagctcaagagctgac 476
Db 295 CCGGAGAACCCGAGCCCTCTGCGCCGCCGCTACATGAACCTTGACCTCAGCGACTTGCC 236
Oy 477 ctctccatcgtttgagaagcgtccatgatgaataatctccactgctgaacccctcctcagcc 536
Db 235 CTCTCATCGTTTGGAGAACGCTCATGATGAATATCTCAACTGCTGAACTCTCTTCAAGCC 176
Oy 537 cactctgtgacaagcttctagttagaagacgtcttgataagtgatgaggaaggaact 596
Db 175 CACTCTGTGGGACAAAGCTTTCTAGTAGACACTCTTGGATTAAGTGCATGAGAGAGAACT 116
Oy 597 gttgaacaattgagaacagaaacccgagattgctg 628
Db 115 GTTGACAATTTGAAGACAGAAACCGGCTAGGTG 84

```

RESULT 9

AA154039/c
ID AA154039 standard; DNA: 499 BP.

AC AA154039;

DT 17-OCT-2001 (first entry)

DE Probe #22725 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;

KM genetic disorder; ss.

OS Homo sapiens.

XX WO200157272-AZ.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0633366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

XX Claim 25; SEQ ID NO 22725; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENPs).

CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.

SO Sequence 499 BP: 116 A; 142 C; 126 G; 115 T; 0 other;

Query Match 13.5%; Score 455.2; DB 22; Length 499;

Best Local Similarly 99.3%; Pred. No. 4.2e-103;

Matches 457; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 169 atgtcgaatggatattccacagagaattccgcatactcactgtgcttcaaggcc 228

Db 499 ATGTGAAATGATTTCCACAGAGAGATTTCCGCTATCTCATCTGTGCTTCAAGGCC 440

Oy 229 aggttgaanaatgatacattcagagatgagcctgtgtgactgaactgaacttctgctga 288

Db 439 AGGGTGAATAATGTACATCCAGCTGAGAGCTGTGCTGAGTACCTGACCTTTGCTGCA 380

Oy 289 gaggatgaagagagagatcagaagagacagtcgcacctccgggaacatgacagcagltgaa 348

Db 379 GAGGTGAAGGAGCAGATTCAGAGGACAGTCCACCTCCGGGAACATGACAGGCGTGA 320

Oy 349 ctgtctgtgagcactctggaagaagagagctgacactgttgagactcgggaattgtg 408

Db 319 CTGTCTGTGAGCACCCTTGAGAGAGGAGTGTGACCTTGATTGGATCGGGAATTTGTG 260

Oy 409 gaggccctccggaagacccgagcctctgcccgcgcctacatgaacccgagctcag 468

Db 259 GAGGCCCTCCGAGAACCCGAGCCCTGTGCGCCGCCGCTACATGAACCTGACTACG 200

Oy 469 gaattgcccctcactcagtttgagaacgctcagatgaataatcacaactgtgaaccc 528

Db 199 GACTTGCCCTCTCCATGCTTTGAGAACGCTCATGATGAATATCTCAACTGCTGAACCTC 140

Oy 529 cttaagccactctggtgagacacctcctagttagaagcgttgtagatgagtgag 588

Db 139 CTTGAGCCCACTGTGTGGACAACTTCTAGTAGAACGCTTGATGAATGATGATGATGAG 80

Oy 589 gaggaaactgtgacaattgagaacagaaacccgagattgctg 628

Db 79 GAGGAACTGTGACATTTGAAGACAGAAACCGGCTAGGTG 40

DE Nucleotide sequence of a human OLD-139 polynucleotide.

KM OLD-35; OLD-64; OLD-137; OLD-139; OLD-142; OLD-175; cancerous phenotype;

KM cellular senescence; terminal differentiation; growth suppression;

KM aging process; type I interferon; cancer cell; tissue regeneration; ss.

OS Homo sapiens.

XX WO200046391-A2.

XX 10-AUG-2000.

XX 02-FEB-2000; 2000WO-US02920.

XX 02-FEB-1999; 99US-0243277.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Fisher PB, Leszczynska M;

Db 308 ccgaattcaaggaacttaccatgttaagaaacaaagacactgccaagaagtgtgc 367
QY 3024 cgactatcaataatgttgaat 3047
|||||
Db 368 cgactatcacaataatgttgaat 391

RESULT 12

AAV89625
ID AAV89625 standard; cDNA: 374 BP.

AAV89625:

15-FEB-1999 (first entry)

EST clone C0522.

Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
gene therapy; ss.

Homo sapiens.

W09845436-A2.

15-OCT-1998.

10-APR-1998; 98WO-US06955.

10-APR-1997; 97US-0838821.

(GENE) GENETICS INST INC.

Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Racie JA, Spaulding V, Treacy M;

WPI; 1999-070077/06.

New polynucleotides encoding human secreted proteins - derived from
e.g. human blood, kidney, foetal lung, placenta, testes, brain,
ovary, pituitary, retina and colon cDNA libraries.

Claim 1: Page 277; 618pp; English.

The present sequence represents a human expressed sequence tag (EST).
The polynucleotide, which is a secreted EST, and the encoded protein
are predicted to have useful biological activities which would make
them suitable for treating, preventing or ameliorating medical
conditions in humans and animals, although no supporting data is
given. Suggested activities include nutritional activity, immune
stimulating or suppressing activity, haematopoiesis regulating
activity, tissue growth activity, activin/inhibin activity,
chemotactic/chemokinetic activity, haemostatic and thrombolytic
activity, receptor/ligand activity, anti-inflammatory activity,
cadherin/tumour invasion suppressor activity, tumour inhibition
activity. The polynucleotide may also be useful for gene therapy.

Sequence 374 BP; 146 A; 47 C; 88 G; 93 T; 0 other;

Query Match 10.9%; Score 365.8; DB 20; Length 374;

Best Local Similarity 99.5%; Freq. No. 5.4e-81;

Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1824 gaaacttcagaataatbacaagaattcaactatgtcaatgagtcacatgca 1883

QY 1884 gaaacttcagaataatbacaagaattcaactatgtcaatgagtcacatgca 1943

Db 62 ttgtgaactcaaccctatgacaatgagccattcaatgagaaataaagctgcaataaga 121

QY 1944 aggaatctgcaagaacgctgtttgtcagaacatttgaagaatgacatgagccctaca 2003
|||||
Db 122 aggaatctgcaagaacgctgtttgtcagaacatttgaagaatgacatgagccctaca 181
QY 2004 aattaatgacacatttcgaatgatagtcgtataccatcttgaacatttcataatga 2063
|||||
Db 182 aattaatgacacatttcgaatgatagtcgtataccatcttgaacatttcataatga 241
QY 2064 aggaagaagataagaagcttgcagtcacalagaagaatgatagtatgagagtgatgata 2123
|||||
Db 242 aggaagaagataagaagcttgcagtcacalagaagaatgatagtatgagagtgatgata 301
QY 2124 gtaattgtatgtatgaataagatgagatgatttaagaacatttgaacatgataaac 2183
|||||
Db 302 gtaattgtatgtatgaataagatgagatgatttaagaacatttgaacatgataaac 361
QY 2184 agatagact 2192
|||||
Db 362 agatagact 370

RESULT 13

AAF66499
ID AAF66499 standard; cDNA: 387 BP.

AAF66499:

09-APR-2001 (first entry)

Novel human polynucleotide, SEQ ID NO: 2255.

Human; cytostatic; gene therapy; colon cancer; prostate cancer;

breast cancer; lung cancer; cancer detection; ss.

Homo sapiens.

W0200102568-A2.

11-JAN-2001.

30-JUN-2000; 2000WO-US18374.

02-JUL-1999; 99US-0142310.

02-JUL-1999; 99US-0142311.

(CHIR) CHIRON CORP.

Williams IT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;

Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;

Chenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;

Kita D, Garcia V, Jones LW, Strache-Grain B;

WPI; 2001-091805/10.

Claim 9; Page 874; 1046pp; English.

The present sequence is one of 3351 sequences in a library of human
polynucleotides. The library is used to detect differentially expressed
genes correlated with a cancerous state of a mammalian cell and can
detect colon, prostate, breast and lung cancer. The library can be used
to produce probes for detection of mRNA and to produce additional copies
of the polynucleotides. The probes can be used for chromosome mapping of
the polynucleotide and for detection of transcription levels. Ribozymes
or antisense oligonucleotides can be generated. The polynucleotides and
their gene products are used as genetic or biochemical markers (e.g. in
blood or tissues) that will detect the earliest changes along the
carcinogenesis pathway and/or monitor the efficacy of therapies and
preventive interventions. The polynucleotides, polypeptides and

CC antihodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 387 BP; 143 A; 51 C; 91 G; 102 T; 0 other;

Query Match 10.6%; Score 356.8; DB 22; Length 387;
Best Local Similarity 96.8%; Pred. No. 9.4e-79;
Matches 364; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

YY 1882 gatttgaactcaacctatgaacatgagccattcaatggaataaaagctgcaaa 1941
|||||
DB 11 gatttgaactcaacctatgaacatgagccattcaatggaataaaagctgcaaa 70

YY 1942 aaaggaatcgaacaaacgctgttctgcaacacattggaagatgacatgagcccta 2001
|||||
DB 71 gaaggaatcgcagagaacgctgttctgcaacacattggaagatgacatgagcccta 130

YY 2002 caaatatgacacaaatcgatgatagcgtatcactccttgaaacttctataat 2061
|||||
DB 131 caaatatgacacaaatcgatgatagcgtatcactccttgaaacttctataat 190

YY 2062 gaaggaatgataaaatttgcagtcataagatgatagtatgaggtgagatgat 2121
|||||
DB 191 gaaggaatgataaaatttgcagtcataagatgatagtatgaggtgagatgat 250

YY 2122 gattatctgattgagatgagatgagatgatttaagaacattgaacttgatgaa 2181
|||||
DB 251 gattatctgattgagatgagatgagatgatttaagaacattgaacttgatgaa 310

YY 2182 acagatgattctcactgacttattttgaaaaaataaattgtgaaaagctggtc 2241
|||||
DB 311 acagatgattctcactgacttattttgaaaaaataaattgtgaaaagctggtc 370

YY 2242 gaaaccccaaatatg 2257
|||||
DB 371 gaaaccccaaatatg 386

RESULT 14

AA090050 ID AA090050 standard; cDNA: 301 BP.

AC AA090050:
DT 27-OCT-1995 (first entry)

XX Melanoma differentiation associated gene mda-5 fragment.

KW Melanoma differentiation associated gene; mda; cancer;
HO-1 melanoma cell; ss.

OS Homo sapiens.

PN W09511986-A.

PD 04-MAY-1995.

PF 24-OCT-1994; 94MO-US12160.

PR 30-SEP-1994; 94US-0316537.

PR 27-OCT-1993; 93US-0143576.

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Fisher PB, Jiang H;

DR WPI; 1995-178878/23.

PT New method for generation of subtracted cDNA libraries - used
PT partic. for identifying melanoma differentiation associated genes
PT for cancer study, diagnosis and therapy.

XX Claim 46; Page 216; 421pp; English.

PS Mda-5 cDNA is novel. It has sequence homology to a Homo sapiens
XX putatively transcribed partial sequence; Accession no. AA220545; from
CC the UK-HGMP, MRC Human Genome Mapping Project Resource, Centre
CC Clinical Research Centre, Watford Road, Harrow, Middlesex, HA1,
CC England. It represents a novel IFN-gamma inducible gene which
CC also displays increased expression during terminal cell diffn.
CC in HO-1 cells. It represents a cytokine-, and terminal
CC differentiation-inducible gene displaying increased expression in
CC all melanomas, in select carcinomas, in normal skin fibroblasts and
CC in both normal cerebellum and GBM.

XX Sequence 301 BP; 110 A; 61 C; 56 G; 74 T; 0 other;

Query Match 8.5%; Score 285.4; DB 16; Length 301;
Best Local Similarity 99.3%; Pred. No. 4.2e-61;
Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

YY 1592 acaaccagatgattcccttccctcagatgactgggaactaacagcttcactggttgaag 1651
|||||
DB 1 acaaccagatgattcccttccctcagatgactgggaactaacagcttcactggttgaag 60

YY 1652 gggccaggaagcaagccaaagctgaaagacacattttaaactatgtgccaatcttgatg 1711
|||||
DB 61 gggccaggaagcaagccaaagctgaaagacacattttaaactatgtgccaatcttgatg 120

YY 1712 cattactatlaaactgtttaaagaacacttgatcacttgaataaaccaatcagagac 1771
|||||
DB 121 cattactatlaaactgtttaaagaacacttgatcacttgaataaaccaatcagagac 179

YY 1772 catcagaagatttgccttgcagatgcacacagagaagatccatttaagaagaacttc 1831
|||||
DB 180 catcagaagatttgccttgcagatgcacacagagaagatccatttaagaagaacttc 239

YY 1832 tagaataatgacagaagatccaacttatttgcanaatgagtcacatcttgga 1890
|||||
DB 240 tagaataatgacagaagatccaacttatttgcanaatgagtcacatcttgga 298

RESULT 15

AA11913 ID AA11913 standard; DNA: 292 BP.

AC AA11913:

DT 12-OCT-2001 (first entry)

DE Probe #9846 for gene expression analysis in human cervical cell sample.

KW Probe: human; microarray; gene expression; cervical epithelial cell;

KM cervical cancer; ss.

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX

PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX

PS Claim 25: SEQ ID No 9846; 487pp; English.
XX

CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 292 BP: 108 A: 31 C: 71 G: 82 T: 0 other:

Query Match 8.4%; Score 281.4; DB 22; Length 292;
Best Local Similarity 97.9%; Pred. No. 4.1e-60;
Matches 285; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1934 ctgcaaaaaaaggaatcgcaagaacgtgttgcagaacatttgaggaagtaaatg 1993
DB 1 ctgcaaaaaaaggaatcgcaagaacgtgttgcagaacatttgaggaagtaaatg 60
QY 1994 aggcctacaattgaacacattgagatagatgagatgacatttgaacttgaactt 2053
DB 61 aggcctacaattgaacacattgagatagatgagatgacatttgaacttgaactt 120
QY 2054 tctataatgaagaagaatlaaagaatttcagtcataagaatgataatgagagtg 2113
DB 121 tctataatgaagaagaatlaaagaatttcagtcataagaatgataatgagagtg 180
QY 2114 gtgatgatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 2173
DB 181 gtgatgatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 240
QY 2174 tggatgaacacagatagatttctcagacttatttttgaacacataaa 2224
DB 241 tggatgaacacagatagatttctcagacttatttttgaacacataaa 291

Search completed: January 17, 2002, 03:26:40
Job time: 7664 sec



...

.

JOURNAL. Submitted (10-SEP-1998) Urology, Columbia University, P&S, 6300
W168th, New York, NY 10032, USA
FEATURES Location/Qualifiers
SOURCE 1. .3380

OY	1561	aagltgaaaaacaatagactcacaagaagaacaacacagtgatctcccttcctcagata	1620
Db	1561	AAGTGTAAAAAACAATAGACTCAAGAAAGAAACAAACAGTGATTCCTCTCCAGATA	1620
OY	1621	ctgagactcaacagcttbaacttggtgtgtgaagggccacagaacaaagccaaagctgaagaa	1660
Db	1621	CTGGGACTTAACAGCTTACCCTGGTGTGTGAGGGGCCACAGAACCAAGCCGTAAGCAA	1660
OY	1661	cacatcttlaaactatgtgcgaactctgtgatcttactatlaaacgttlaaagaaac	1740
Db	1661	CACATTTAAAAACTATGTGCCAATCTTGATGCATTTACTATTAACAGCTTTAAAGAAAC	1740
OY	1741	cttgatcaactgtaaaaccaacaatacagagcgcatgcagaagtltgccattgcagatga	1800
Db	1741	CTTGATCAACTGTAAAAACCAAAATACAGGAGCGCATGCAGCAAGTTTGCCATTGCAGTGA	1800
OY	1801	accagaagaatccattlaaagaagaactctagaataatagacagagattcaactat	1860
Db	1801	ACCGAGCAAGATCCATTAAAGAGAACTCTAGAAATATATACAAAGATTCAAACTTAT	1860
OY	1861	tgtaaatgagtcceaatgtcagaatlttggaactcaacctgaacaaatggccattcaa	1920
Db	1861	TGTCAATGAGTCCCAATGTAGATTTTGGAGCTCAACCTATGACAACTGGGCCATTGAA	1920
OY	1921	atggaaaaaaangctgcgaaaaaaaggaatatcgcaagaacgtgtgtgtgcgaacattg	1980
Db	1921	ATGCAAAAAAANGCTGCANAAAAAGGAATAATCGCAAGAACGTGTTGTGCAGAACATTG	1980
OY	1981	aggaagatacaatgagggccctacaaatlaatagacacaaatcgaaatgatagctatact	2040
Db	1981	AGGAAGATCAATGAGGCCCTACAAATTAATGACACAAATTCGATATATGATGATCGTATCT	2040
OY	2041	catcttgaactcttctaatagaagaagaagaataaagaatctgcagatataagaatgat	2100
Db	2041	CATCTTGAACCTTCTATATATGAAAGAAAGATAAACAATTTGCAGTCTCAAGAAATAT	2100
OY	2101	agtgatgaaggtgtgtgatgatgatgtgtgatgtgtatgaagaatgagatgatataag	2160
Db	2101	AGTGATGAGGGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATTAAG	2160
OY	2161	aaacctttgaacagtgtgaagaacagatagatcttctatacttcttttttgaaacacat	2220
Db	2161	AAACCTTTGAACAGTGATGTAACAGATAGATTTCTCATACTTTATTTTTCGAAAAACAT	2220
OY	2221	aaaaatgttgaaaagcgtgcgtgtgaaaacccagaaatagaaatgaaagctgacaaata	2280
Db	2221	AAAAATGTTGAAAAGCTGCGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTA	2280
OY	2281	agaaataccataatgtgagcaataactactagactgaagaaatcagcacagagaataacttc	2340
Db	2281	AGAAATACCAATATGAGCAATATACTAGACTGAGGAGATCGACGAGGAATAATCTTT	2340
OY	2341	acaaaaacacagacagatgtcatatgtgccttcccaatgtgatatctgaaatatgaaaaatt	2400
Db	2341	ACAAAAACACGACAGACTGTGATATGCGCTTTCCCATGTGATTACTGAAAATGTAAAAATT	2400
OY	2401	gctgaagtaggaagtcagaagcccaacatctgattggagctgcgagacagcagtgaaftcaa	2460
Db	2401	GCTGAGTAGGAGTCAAGGCCCAACCATCTGATTGGAGCTGGACACAGCGTAGATTCAA	2460
OY	2461	cccatgacacagaaatgaaacaaaagaagtcattagtaaatcttcgcacttggaaaaatcat	2520
Db	2461	CCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAATTTCCGACTGSAAAATCAAT	2520
OY	2521	ctgcttatacgttaccacaagtggcagaagaagctctgatatataaagaatgfaaacatgtc	2580
Db	2521	CTGCTTATCGCTTACCACAGTGGCAGGAAGAGCTGTGATATTAAAGAAATGTAACTTGT	2580
OY	2581	atccgttctatgtctcgtcaccaatgaaatgccaatgtctcaagcccggtgtgtgaacaga	2640
Db	2581	ATCCGTTATGTGCTCTGTCACCAATAGAAATAGCCATGTCTCAGGCCCGGTGTGTGACCCGA	2640
OY	2641	gctgatgagagcacctacgtctcgtgtgtgtcacaagtgattcagagaattatgcaacatgag	2700

Db	2641	GCATGTAGACACACTACAGTCCTCTGGTGTCTACAGTGGTTCCAGAGTTATCGAACAATGAG	2700
QY	2701	acaagtaatgattcccgagagaagatgatgataaagctatactatgcttcaaaatg	2760
Db	2701	ACAGTTAATGATTTCCGAGAGAACATGATGTATATAACCTATACATTGTGTTCAAAATATG	2760
QY	2761	aaaccagaagagatgctcatalaagatttggaaatlacagatgcaagatataatgqaaag	2820
Db	2761	AAACCAGAGSAGATATGCTCATACATTTTGGATTACAGATGCAAAATATATGSAANAAG	2820
QY	2821	aaaatgaaaccagaagagaatatatgccaagatatacaagaataaaccataactaact	2880
Db	2821	AAAATGAAACCAACAGAAATATTGTCCAAAGCATTCAGAAATAACCATCATCTAATAACT	2880
QY	2881	ttcccttgcagaanaactgagatgctgctagccgtctctggggaagatatcagtgtaattgag	2940
Db	2881	TTCCCTTTCCAAAACACTGCAGTGTGTCTACCCCTGTTCTGGGAGAGATATCCATGTAATTGAG	2940
QY	2941	aaaatgcatcagctcaatataatgaccccaagaattcaagaagacttcaatigtlaagagaaac	3000
Db	2941	AAAATGCATCAGCTCATATATGACCCCAAGAAATTCAAGSAACTTTCATTTGTAAGGAAAC	3000
QY	3001	aaagcactgcaaaagaagtgctgcccaactatacaataaattggtgaatcactgcaaatgt	3060
Db	3001	AAAGCACTGCAAAAGAAAGACTGTGCCGAGTATCAATAATATGTTGTAATCATCTGTCAAAATGT	3060
QY	3061	gcccagagcttggggagacaatgatgtagcaagaagaagcttaagattgacctctcaaaata	3120
Db	3061	GGCCAGGCTTGGGGCAACATGATGTGTGCACAAGGCTTAGTTGGCTTGCTCTCAAAATA	3120
QY	3121	aggaatttltgtatgtgttttcaaaataatcacaacaaagaacaatacaaaagtggtga	3180
Db	3121	AGGAATTTTGTATGTGTGTTTCAAAATATATTCACAACAAAGAAACATTCAAAAGTGGGTA	3180
QY	3181	gaattacctatacaatttcccaactggaactatcagaatgcgtgttattatgtatgtgag	3240
Db	3181	GAATTACCTATCACATTTTCCCAATCTTGCACTATTCAGATCTGTTTATTTAGTATGAGAG	3240
QY	3241	gattagacgttattggaagatcttctttaaataactatcagaatgaacatttaaatgattga	3300
Db	3241	GATTAGCACTTGATTGGAAGATCTTTTAAATTACTATCAGTTTAACATTTAATATGATTA	3300
QY	3301	tgattaatgtatcattatgctacagaaactacataagaatacaataaattgattgttla	3360
Db	3301	TGATTAAATGATATCTATTATGCTACAGAACTACATTAAGAAATCATTAAATGATTGTTTA	3360
QY	3361	cctctg 3365	
Db	3361	CTCTG 3365	
RESULT 2			
LOCUS	AY017378	3373 bp mRNA	PRI 31-JAN-2001
DEFINITION	Homo sapiens RNA helicase-DEAD box protein RH16 mRNA, complete cds.		
ACCESSION	AY017378		
VERSION	AY017378.1	GI:12621065	
KEYWORDS	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 3373) Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.		
TITLE	Identification of a new RNA helicase (RH16) regulated by the immunomodulator Murabutide		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3373) Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.		
AUTHORS			

OY	1515	caagaagacgctgtataataaactcctatggcgattatatttgatgcgaagtttgaaaaaaca	1574
Db	1501	CAAAAGAAAGCGTGTATATATATACATCATGAGGCGATTATTGATGCGAAGATTGAAAAACA	1560
OY	1575	taagctcaagaagaanaacaaccatgattccctctccagataactggactaaagc	1634
Db	1561	TAGACTCAAGAAAMAAACCAACCACTGATTTCCCTCTCGATACTGGGCTTACAGC	1620
OY	1635	ttcaactgtgtttgaggggcccagaagccaaagctgaagaacacatttlaaact	1694
Db	1621	TTCACTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACATTTTAAACT	1680
OY	1785	aaaccaatcaagaagcactgacaagaagtttccatctgacaagtgaaaccaagaagctc	1814
Db	1741	AAACCAATACAGAGCGCATGCAGAAAGTTTGGCATTTGCAGTGCACACCAGAAAGATCC	1800
Db	1815	atttaagaagaaactttcagaataatgacaagatlcacaattltgcaaatgagtc	1874
Db	1801	ATTTAAGAGAACTTCTAGCAATATATGCAAGATTTCAACTTATTTGCAATGATGCC	1860
OY	1875	aatgtcagatttttggaactcaaccctatgaacaatgagccatccaatggaaaaaagc	1934
Db	1861	AATGTCAGATTTTGGAACTCAACCCATATACCATGGCCATTCAAATGGAAAAAAGC	1920
OY	1935	tgcaaaaaagaagatcgcagaagaacgttttctgacagaactttgaagaatlaacatga	1994
Db	1921	TGCAAAACGAAGAAATGCGAAACAAAGCTGTTTGTGAGAACCTTTGAGCAACTACAAATA	1980
OY	1995	ggccctcaaatlaatgacacaacttcgaatgatagatgctatctactccttgaacatt	2054
Db	1981	GGCCCTCAAAATTTATATACACCAATTTGGAATGATAGATGCGTATACATCTTGAAACTT	2040
OY	2055	ctataatggaagaagaagaataagaagtttgcagtcataagaagatgattgatgaagctg	2114
Db	2041	CTATATATCAAGCAAGAAATATGAAGATTGTGCAGTCATAGAGATGATAGTAGAGCGTGG	2100
OY	2115	tgatgatgaagatctgtagatgattgaagaatgagatgagatgtttaaagaaccttgaact	2174
Db	2101	TGATGATGAGTATTGTGATGCTGATCAGATGAGATGATTTTAAACAAACCTTTGAAACT	2160
OY	2175	ggatgaacaacagatagattcttcatagacttatttlttgaacaacataaattgttlaaag	2234
Db	2161	GGATGAAACACATGATGATTCTTCATGACTTATTTTGAACAAACATTAATATGTGAAGG	2220
Db	2235	gctgcttgaaaaeccagaataatgaanaatgaaaagctgcaccaatlaagaatatccataat	2294
Db	2221	GCTGCTGTGAAACCCACCAATATGAAATGAAAGCTGACCAAAATTATACAAATATCCATAT	2280
OY	2295	ggaggaatatactagagactgaagaatcagacagaggaataactttacaanaaacgcgca	2354
Db	2281	GGAGGAATATCTAGGACTAGGAATCAGACAGGAGAAATACTTTACAAAAACGCA	2340
OY	2355	gagtgcatatagcctttcccgatggattacttgaanaatgaaaatttgcctgaagatgaggt	2414
Db	2341	GAGTGCATATCGGCTTCCGACAGTGAATACTGAAATGAAAAATTTCGTAAGTAGAGCT	2400
OY	2415	caaaagccacatctgatctgagctctgacacagcagctgactccaacccatgaacagaa	2474
Db	2401	CAAAAGCCCAACCATCTGATTGAGCTGAGCACACAGCACTGATTCMAAACCATGACACAGA	2460
OY	2475	tgaacaaaaaagaatctatgtaaatcttgcacatgtaaaaaatccaatctgattatgctac	2534
Db	2461	TGAACAAAAAAGACTCTATTAGTAATTTCCACTGAGAAAAATAATCTGCTTATGCTTAC	2520
OY	2535	cacagtgcacaaaagaagctctgatatlaagaatgtaacatctgtatccgttatgctct	2594
Db	2521	CACAGTGCACAAACAGAGCTCGGATATTAAAGATATTAACATTGTTATCCGTTATAGGCT	2580
OY	2595	cgtcaaccaatgaatagccatggtccaagccgttgctgaagccagaagctgatgaagacac	2654

Db	261	CGTCAACCAATGAAATATAGCCATGTGTCAGGCCCGCTGTGTGAGCCAGAGCTATATGAGACAC	2640
QY	2655	ctacgtccctgtgtctcacaatgtgttcaagagatattcgaacatgaacatgaacgttaattt	2714
Db	2661	CTACGTCCTCGTGTCTCAGCACTGGTTTCAGGAGTTATGCAAGCTGAGACAGTATATGATT	2700
QY	2715	ccgagagagaagatgtatgataaagctatacatgttgttccaataatgaaccagagagta	2774
Db	2701	CCGAGAGAGAGATGTATGTATAAGCTATACATTGTGTTCAAAATATGAAACAGAGGAGTA	2760
QY	2775	tgtcctataagtttttggtaattacagatgccaagatataatgtgaaaaaagaataaaaccaa	2834
Db	2761	TGCTCATAAACAATTTTGGCAATTTACAGATGCAAAAGATATATGAGAAAGAAATGAAACCAA	2820
QY	2835	gagaatatattccaaagcatcaagaataaacccaataaacttcttttgcacaaa	2894
Db	2821	GAGAAATATTGCCAAGCATTCACAAAGATTAACCCATCATTATAACTTCTTGTGCAAAA	2880
QY	2895	ctcgagtgctgaagctgtgtctcgaggagaatataccatgtaattgagaaaaatgcatca	2954
Db	2881	CTGCAGTGTGCTTACCTGCTTGTGGGGAAGATATCCATGTAATTGAGAAATGCAATCAGT	2940
QY	2955	caatatgacccccagaaattccaagaaactttacattgttaagagaacaaagacatcgcaaa	3014
Db	2941	CAATATGACCCCGAATTTCAAGGAATTTACATTGTAAAGAGAAACCAACCACTGCAAAA	3000
QY	3015	gaagtgtccgaactatcaataaattgtgaaatcatctcgcaatgtgtgccaggtttgag	3074
Db	3001	GAACTGTGCCCATATCATCAATTAATATGTAATATATCTGCAATATGTGGCCAGGCTTGCGG	3060
QY	3075	aacaatgatgtgcacaaaggcttagatttgccttgccttgcacaaataagaaatttgtagt	3134
Db	3061	AACAATGATGTGTGCAACAAGGCTTAGATTGCTTGTCTCAAAATAGGAATTTGTAGT	3120
QY	3135	ggttttcaaaaataattcaacaagaacaaatcacataaaaaatggtggttaaatlactatcac	3194
Db	3121	GGTTTCAAAAATATATTCACAAAGAAACAAATACAAAAGGGGTACAAATTTGCTATCAC	3180
QY	3195	atttcccaacttgaactatcagaatgctgttattttagtataagattgaacacttgat	3254
Db	3181	ATTTCOCATCTTGCACTATTGCAATGCTGTTTATTAGTATGAGCATTAAGCACTTGAT	3240
QY	3255	tgaagatctcttllaasaalactatcaagttaaacacttlaatatgattatgattatgattc	3314
Db	3241	TGAACATCTCTTAAATACTATCATGTTAAACATTTAATGATTGATTAATGATATTC	3300
QY	3315	attatgtacagaaactgacataaagaatacaataaataatgattgtcttactct	3364
Db	3301	ATTATGTACAGAACTGACATTAAGAAATCAATTAATTAATGATTTGTTTACTCT	3350
RESULT 3			
LOCUS	BC007966	1213 bp mRNA	PRI
DEFINITION	Homo sapiens, clone IMAGE:4304805, mRNA.		12-JUL-2001
ACCESSION	BC007966		
VERSION	BC007966.1	GI:14044090	
KEYWORDS	human.		
ORGANISM	Homo sapiens		
REFERENCE	Enkavola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1213)		
TITLE	Strausberg,R.		
JOURNAL	Direct Submission		
COMMENT	Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
CONTACT	Contact: MGC help desk		
EMAIL	Email: cgapops-rt@mail.nih.gov		

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Sequencing Center (NISC),
 Gaithersburg, Maryland.

Web site: <http://www.nisc.nih.gov/nisc-mgenhgrl.nih.gov>
 Contact: Shevchenko, Y., Wetherby, K.D., Bouckstrom, Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stancirpop, S., Thomas, P.J., Tlonson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IMAGE Plate: 20 Row: P Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source location/Qualifiers
 1..1213
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4304805"
 /tissue_type="uterus, leiomyosarcoma"
 /clone_lib="NIH MGC_46"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 BASE COUNT 386 a 247 c 285 g 295 t
 ORIGIN

Query Match 35.5% Score 1194.4; DB 9; Length 1213;
 Best Local Similarity 99.5%; Pred. No. 3e-242;
 Matches 1198; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 380 ggcaccttggatggaactcgaatcgtggaagccctcgaagaacgagcagcctctg 439
    |||
DB 9 ggcaccttggatggaactcgaatcgtggaagccctcgaagaacgagcagcctctg 68
    |||

QY 440 ccgcccgcatacgaacccctgagcagcagcctcgcctccatcgtttgagaagcctc 499
    |||
DB 69 ccgcccgcatacgaacccctgagcagcagcctcgcctccatcgtttgagaagcctc 128
    |||

QY 500 atgataatcctcgaacccctgagcagcagcctcgcctccatcgtttgagaagcctc 559
    |||
DB 129 atgataatcctcgaacccctgagcagcagcctcgcctccatcgtttgagaagcctc 188
    |||

QY 560 ttgaagacgcttggataagtcgaaggaaggaacgttgacaattgaagacagaac 619
    |||
DB 189 ttgaagacgcttggataagtcgaaggaaggaacgttgacaattgaagacagaac 248
    |||

QY 620 ggaattgctgctgagcaagaacaatgaatgaatcagtgatgaagagagctactaaaga 679
    |||
DB 249 ggaattgctgctgagcaagaacaatgaatgaatcagtgatgaagagagctactaaaga 308
    |||

QY 680 ttgtcagaagaagaacgagcttctgcatcttgaatgcttctgttaaacagaaga 739
    |||
DB 309 ttgtcagaagaagaacgagcttctgcatcttgaatgcttctgttaaacagaaga 368
    |||

QY 740 atgaacttgcacaagatgaacagcctctgattgctcagaagaacatgaagattgaga 799
    |||
DB 369 atgaacttgcacaagatgaacagcctctgattgctcagaagaacatgaagattgaga 428
    |||

QY 800 attatcacaagtgatgctcctcagaagagagcaactcttcaaccacagcttcagc 859
    |||
DB 429 attatcacaagtgatgctcctcagaagagagcaactcttcaaccacagcttcagc 488
    |||

QY 860 caaatctgagagagagctctgaggaatgaggaataactcaatcaatcattcttggag 919
    |||

```

```

DB 489 CAATCTGGAGAGGAGCTGTGGGGCATGGAGAAATAACTCATCGAATCATTTTGGCAG 548
QY 920 attctctgagatcccaatcagacacaagtctgagagaagagtgatcagctgtag 979
    |||
DB 549 ATTCTTCTGTGATGTTTCAATATCAGACACACAACTTTGGAGAAAGAGTGTAGCTGCTTAG 608
    |||

QY 980 atgaagctcttggacataacagcaacatgagcagtgatctcagcaccatggaagatgatt 1039
    |||
DB 609 ATGAAGAGCTTGGACATTAACAGCAACATGGGCACTGATTCAGGACCATGGAGATGATT 668
    |||

QY 1040 cagatgagaagaatgagtgagcagaagacatcccgagagcagcaatcctcagctcagctt 1099
    |||
DB 669 CAGATGAAGAGAAATGTGACAGCAACACATCCCGAGGCGCAGAGCTCAGCTAGGCTT 728
    |||

QY 1100 acccaatgagaagtgagcagcagccttgagaaggaatatacatatctgctcccta 1159
    |||
DB 729 ACCAAATGGAAGTGGCCAGCAGCAGCTTGGAGGAGAAATATATATCTGCTCCCTTA 788
    |||

QY 1160 caaggaatggaagaaacccaagagtgatcttcaatctgccaagatcactagacaagaaga 1219
    |||
DB 789 CAGGAGTGCAGAAACAGAGAGTGTGATTCATTCAGCAAGATCCTTAGACAGAGAGA 848
    |||

QY 1220 aaaaagatctgagcctcgaagaagtatagtcttctgcaataagatcagctgattgaa 1279
    |||
DB 849 AAAAAGATCTGAGCCTGGAGAAAGTTAGTCTTGTCAATTAAGCTACTGCTAGTTGAA 908
    |||

QY 1280 agcctctcgaagagatctcgaacatcttgaagaagatgatacgtgattatgattaa 1339
    |||
DB 909 AGCTCTTCCGAGAGAGTGTCAACCATTTTGAAGAAATGTATCTGTTATTGGATTAA 968
    |||

QY 1340 gtgagtaacccaactgaataatcattcagaagtgatgagctgagctgataattaa 1399
    |||
DB 969 GTGCTGATACCCAGCACTGAATAATATCATTTCCAGAAAGTTGCAAGTCTGATATATTA 1028
    |||

QY 1400 tcaatcagcctcaatcttgaagaactccctcttaacttgaagaatgaggaagatgctg 1459
    |||
DB 1029 TCACTACAGCTCAAAATCTTGAAGATCCCTTTAACTTGAAGAAATGGAAGAAATGCTG 1088
    |||

QY 1460 gtgtcaatgtgcagacttctccctcattatcattgataatgcatcaccaacaaga 1519
    |||
DB 1089 GTGTCAATGTGACAGCTTTTCCCTCATATGATGATGATGATGATGATGATGATGATG 1148
    |||

QY 1520 aagcagtgataataacatcagcagcagcatttgaatgaggaatggaagaacatagac 1579
    |||
DB 1149 AAGCAGGTATATATATATATATGATGATGATGATGATGATGATGATGATGATGATG 1208
    |||

QY 1580 tcaa 1583
    |||
DB 1209 AAAA 1212
    |||

RESULT 4
BC004031 2050 bp mRNA KOD 12-JUL-2001
LOCUS BC004031 Mus musculus, clone IMAGE:3495361, mRNA, partial cds.
DEFINITION BC004031
ACCESSION BC004031
VERSION BC004031.1 GI:13278455
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2050)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.

```



```

||||| 1438 TTCAAAATTAACCTGACGCAAGCAAGATGAGTGGGGAATTCCTATAGATTT 1497
||||| 3199 cccaacttgatcattcagctgttattatgaagatgagatgacattgaa 3258
||||| 1498 CCGATCTTGAATCTGCAATAGTCTTGTATGATGATGAGATTAGACTTGATTTAT 1557
||||| 3259 gattctttaaataactacagttaaacattaaatagattatgattatcattaa 3318
||||| 1558 GATTATTTAAATTTGCTCAACTTCACATGATGATGATTTATTTGTTTGGCTA 1617
||||| 3319 tgcatacagaactgacataagaatcaataa 3348
||||| 1618 CACTACTGAGCTAACTTGAATATATGTAGA 1647

RESULT 5
LOCUS AX098232 1258 bp DNA PAT 30-MAR-2001
DEFINITION Sequence 7 from Patent WO0118208.
VERSION AX098232
KEYWORDS AX098232.1 GI:13515364
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1258)
AUTHORS Peyman,J.A., da Silva,A., Hochman,P. and Hsu,A.
TITLE Interferon induced polynucleotides and proteins encoded thereby
JOURNAL Patent: WO 0118208-A 7 15-MAR-2001;
Curagen Corporation (US) ; BIOGEN, INC. (US)
LOCATION/Qualifiers
1..1258
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..915
/note="unnamed protein product"
CDS
/protein_id="CAC35440.1"
/db_xref="GI:13515365"
/translation="MVSNTTPRDVIKSPVAOSGRPCSKMAADSGTHISVPOT
APSAMENDPLSLRKDIYEVFRRPEPLSHLVIOGGENVHPEVKEDIORIVATSGNM
OAVELLSTLEKGVNHLGWTREYALRKRTGSPPLAARYMNPETDLPSPSENAHDEY
LQILNLOPTLVKLLVRIYDLKCEBELITIEDNRNIAAENNGNSGVRELLKRLV
OKENMFAPLNVLRQTNNEIYVELTSGDSSESNATEINLSQVDVDPVEOLLSTTVQ
PNLEKVENFSAFLNVLRTGNNEIYVELTSGDSSESNATEINLSQVDVDPVEOLLSTTVQ
BASE COUNT 399 a 267 c 276 g 316 t
GIN
Query Match 23.9%; Score 803; DB 6; Length 1258;
Best Local Similarity 95.5%; Pred. No. 1.7e-159;
Matches 853; Conservative 0; Mismatches 0; Indels 40; Gaps 1;

```

```

Db 257 -CAGAGGTGAAGAGGAGATTCAGAGACAGTCCGACCTCCGGGACATGAGGAGT 314
||| 345 tgaactgctgagacaccttgagaaagagatctggacacttgatlgagactcggaatt 404
||| 315 TGAAGTCTGCTGAGACACCTTGGAGAGGAGCTGTGGCACCTTGTTGGACTGGGAATT 374
||| 405 cgtggaagccctccggaacccggaacccctctgagcccgatcatalgaacccgagct 464
||| 375 CGTGAAGGCCCTCGGAGAACCGGACACCTCTGGCGCCGCTGACATGAACTGAGCT 434
||| 465 cagagacttccctctccatcgtttgagaaagctcatalgaatgaaatccaaactgtgaa 524
||| 435 CACGACTTGGCTCTCATGCTGTGAGAACGCTCATGATGAAATCTCCAACTGGTGA 494
||| 525 cctcctcagccacactctgtagaagaacttctlaqilaagaagacttgat aatgtcat 584
||| 495 CTTCTTCAAGCCACCTGTGTGGACAGCTTCTAGTAGAGAGCTCTGGATTAAGTGCAT 554
||| 585 ggaaggaagaaactgtgaaacttgagaaagaaacacggaattgtctgagaaacaaatg 644
||| 555 GAGAGGAGAACTGTGCAATTTGAGAGACAGAAACCGGATTGCTGCAGAAAACAAATG 614
||| 645 aaatgaacagagtgtagaagagactaaagagattgtgcaaaaagaaactggtctc 704
||| 615 AAATGAATCAGGTGTAGAGAGCTACTAAAGAGATTGTGCAGAAAGAAATCGTCTTC 674
||| 705 tgcattctgaatgcttctgcaacagaaacaaactgtgtccaaagatlaaacag 784
||| 675 TGCATTTCTGAATTTCTTGTGTCACAGCAACAAATGTTGTCACAGTTAACAG 734
||| 765 ctctgattgctcagaagaactgagagattgagaatttcaaaagttgagtgctca 824
||| 735 CTCTGATTGTCACAAAGCAATGAGAGATTGAAATTTATCAAGATTGATGCTCTCA 794
||| 825 agtgaagaaacacttcttcaacacagcttgcacaaactctgaaagagagctggag 884
||| 795 AGTGAAGAGCACTTTTTCACACAGTTCAACCAATCTGGAGAGAGGCTGGGG 854
||| 885 catgagaataaactcagaatcacttcttgcagacttcttgaatctgagcttca 937
||| 855 CATGGAATTAACATCATCATCATCTTTGCAGATTCTTGTAATTTCAG 907

RESULT 6
LOCUS AX098236 1270 bp DNA PAT 30-MAR-2001
DEFINITION Sequence 11 from Patent WO0118208.
ACCESSION AX098236
VERSION AX098236.1 GI:13515371
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1270)
AUTHORS Peyman,J.A., da Silva,A., Hochman,P. and Hsu,A.
TITLE Interferon induced polynucleotides and proteins encoded thereby
JOURNAL Patent: WO 0118208-A 11 15-MAR-2001;
Curagen Corporation (US) ; BIOGEN, INC. (US)
LOCATION/Qualifiers
1..1270
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..927
/note="unnamed protein product"
CDS
/protein_id="CAC35442.1"
/db_xref="GI:13515372"
/translation="MPAAVNSNTTPRDVIKSPVAOSGRPCSKMAADSGTHISVP
POTAPSAMENDPLSLRKDIYEVFRRPEPLSHLVIOGGENVHPEVKEDIORIVAT
SGNMQAVELLSTLEKGVNHLGWTREYALRKRTGSPPLAARYMNPETDLPSPSENA
HDEYDLNLOPTLVKLLVRIYDLKCEBELITIEDNRNIAAENNGNSGVRELLKRLV
OKENMFSAFLNVLRTGNNEIYVELTSGDSSESNATEINLSQVDVDPVEOLLSTTVQ
KRIVOKENMFSAFLNVLRTGNNEIYVELTSGDSSESNATEINLSQVDVDPVEOLLSTTVQ

```


Db 555 GGAGAGGAGCACTGTTGACAAATTGGAAGACAGAAACCGGATTGCTGCGAGAAACCAATGG 614
 QY 645 aaatgaatcaagtgtaagagagactactaaagaattgtgcagaagaagaactgtgtc 704
 Db 615 AAATGATCAGGTGTAAGAGAGACTACTAAAGAGATGTGCGAGAAAGAAACTGTTCTC 674
 QY 705 tgcattctgaattgtcttcgtcaacaggaagaataaactgtgtcagaaggttaacagg 764
 Db 675 TGCATTCTGAAATGTTCTTCGTCAGAAAGCAATGCACTTGTCCAGAGATTAAACAGG 734
 QY 765 ctctgattgtcacaagaacatgcagagattgagaatttacaagaagttaagtcc 821
 Db 735 CTCTGATTGCTCAGAAAGCAATGCAAGTATTGTAATTTACTGAGAGATTCTTC 791

RESULT 8
 AX098238 3704 bp DNA PAT 30-MAR-2001
 LOCUS AX098238
 DEFINITION Sequence 13 from Patent WO0118208.
 ACCESSION AX098238
 VERSION AX098238.1 GI:13515375
 KEYWORDS
 SOURCE human.
 ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Peyman, J. A., da Silva, A., Hochman, P. and Hsu, A.
 TITLE Interferon induced polynucleotides and proteins encoded thereby
 JOURNAL Patent: WO 0118208-A 13 15-MAR-2001;
 Curagen Corporation (US); BIOGEN, INC. (US)

FEATURES
 source
 1..3704
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..816
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC35443.1"
 /db_xref="GI:13515376"
 /translation="MPAAVNSNTTTPRDVYKSPVAVSGRPRCSKMAADSGTHISPV
 POTAPSAWENPULSLKKKQVEMFHRREFPLSLVIOGGFNNHPRVKQIORTVAT
 SCNMQAVELLSTLEKGVNHLGRTREFVLRKTSPLAATMNPDLTLPSPFENA
 HDEYLOLNLPTLVKLVLDVLDCEMEELTITIDRNRTAAENNGNSGVRELL
 KRIVQENWFSAFLNVLROTGNELVOTLGSDESNAGICNTEEDSSNSA"

BASE COUNT 1168 a 713 c 809 g 1014 t
 ORIGIN

Query Match 19.7% Score 663; DB 6; Length 3704;
 Best local Similarity 92.9%; Pred. No. 7; le-130;
 Matches 722; Conservative 0; Mismatches 15; Indels 40; Gaps 1;

QY 45 gtagcagagcagatgagacacctgtctctcttaagtggcagcagcagcagcagcagcattca 104
 Db 67 GCACGAGCGGTAGACCCCTGCTTCTTAAGTGGCAGCGGACGCGACACCACTTCA 126
 QY 105 cctgtcccgac 164
 Db 127 CCTGTCGCCGAC 186
 QY 165 aaagatgtcgaatgtatctccacacagacagaatttcgcatatcatctcgtgtcattag 224
 Db 187 AAGATGTCGAATGGGTATTCACAGACAGAGATTTCGCTATCATCTGCTGCTTCAG 246
 QY 225 ggcacaggtgaaatgtatcacaccagtgagccgtgtcgtgaactcaactcactctgcgc 284
 Db 247 GCCCGGGGTGAAATGTACATC----- 268
 QY 285 tgcagaggtgaaagagcagatttcagagacagtcgacacctcgagaaacatcagcagcagt 344
 Db 269 -CAGAGGCTGAAGAGCAGCATTCACAGACAGTCCGCCACCTCCGGAACATCAGCAGCAGT 326

QY 345 tgaactcgtctgtgcacacttgagagaaggagtcgtgacaccttggttgacctcggagatt 404
 Db 327 TGAATCTCTGCTGACACCTGTGAGAGGAGAGTCTGCGACCTGTTGAGCTGGGGAATT 386
 QY 405 cgtgaagagccctccgagagaacccggcagacctctggccgcccgttgaactgaacctgact 464
 Db 387 CGTGAGAGCCCTCCGAGAACCGGACCCCTGTGGCCGCCCTGACATGAAACCTTGAGCT 446
 QY 465 caagacttgcctctcctcactcacttggagaacgctcatgatatatctcaactgtgaa 524
 Db 447 CAGGACTTGCCCTCTCATCTGTTGAGAACGCTCATGATGATATATTCGCAACTGCTGAA 506
 QY 525 cctccctcagcccaactcgtgtgagcaagcttctagttgagagcgtcttgataagtcact 584
 Db 507 CCTCTTACGCCCACTGCTGTGGAGCAAGCTTCTAGTTAGAGAGCTTGTGATAGTGCAT 566
 QY 585 ggaagaggaactgttgaacattgagaacagaagaacaggattgtcgtgtgcagaagaactg 644
 Db 567 GGAGGAGCAACTGTTGCAATTTGAGACAGCAAAACCGGATTGCTGCTGCAAAAACAATGG 626
 QY 645 aaatgaatcaagtgtaagagagactactaaagaattgtgcagaagaagaactgtgtc 704
 Db 627 AAATGATCAGGTGTAAGAGAGACTACTAAAGAGATGTGCGAGAAAGAAACTGTTCTC 686
 QY 705 tgcattctgaattgtcttcgtcaacaggaagaataaactgtgtcagaaggttaacagg 764
 Db 687 TGCATTCTGAAATGTTCTTCGTCAGAAAGCAATGCACTTGTCCAGAGATTAAACAGG 746
 QY 765 ctctgattgtcacaagaacatgcagagattgagaatttacaagaagttaagtcc 821
 Db 747 CTCTGATTGCTCAGAAAGCAATGCAAGTATTGTAATTTACTGAGAGATTCTTC 803

RESULT 9
 AC010876/c 95417 bp DNA PRI 30-SEP-2000
 LOCUS AC010876
 DEFINITION Homo sapiens BAC clone RP11-214A4 from 2, complete sequence.
 ACCESSION AC010876
 VERSION AC010876.2 GI:7243901
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM human.

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 TITLE 1 (bases 1 to 95417)
 JOURNAL Sultana, J.F. and Waterston, R.
 TOWARD A COMPLETE HUMAN GENOME SEQUENCE
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 REFERENCE
 AUTHORS 2 (bases 1 to 95417)
 TITLE Waterston, R.H.
 JOURNAL Direct Submission
 Submitted (25-SEP-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 95417)
 REFERENCE
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 Submitted (15-MAR-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 95417)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submission
 Submitted (30-SEP-2000) Department of Genetics, Washington
 University 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 15, 2000 this sequence version replaced g1:5923734.

Center: Washington University Genome Sequencing Center

Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_NH0214A04

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tano, M., Catanesi, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-576116, 200 bp overlap; the clone sequenced to the right is RP11-204024. Actual start of this clone is at base position 126545 of RP11-576116; actual end is at base position 81443 of RP3-404K21.

The sequence RP11-214A4 from base position 23333 to 23415 was derived from PCR off clone DNA.

FEATURES

source
 1. 95417
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-214A4"
 /clone_lib="RPCI-11"
 repeat_region
 248. 269
 /rpt_family="AT-rich"
 repeat_region
 358. 383
 /rpt_family="AT-rich"
 repeat_region
 695. 726
 /rpt_family="AT-rich"
 740. 950
 /note="similar to Mus musculus EST AA123960 (NID:q1682517)
 mp78b04.r1"
 misc_feature
 763. 950
 /note="similar to Mus musculus EST AA656698 (NID:q2592852)
 vp95c09.r1"
 repeat_region
 2364. 2505
 /rpt_family="MERL-type"
 repeat_region
 2714. 2828
 /rpt_family="MIR"
 repeat_region
 2916. 3111
 /rpt_family="MIR"
 repeat_region
 5265. 5370
 /rpt_family="MIR"
 repeat_region
 6509. 6734
 /rpt_family="MIR"

/note="similar to Mus musculus EST AA123960 (NID:q1682517)
 mp78b04.r1"
 misc_feature
 6509. 6710
 /note="similar to Mus musculus EST AA656698 (NID:q2592852)
 vp95c09.r1"
 repeat_region
 7292. 7336
 /rpt_family="(CA)n"
 repeat_region
 7585. 7694
 /rpt_family="AT-rich"
 repeat_region
 8919. 8968
 /rpt_family="L2"
 repeat_region
 10445. 10499
 /rpt_family="L2"
 repeat_region
 11237. 11409
 /rpt_family="L1"
 repeat_region
 11552. 11585
 /rpt_family="L1"
 repeat_region
 11938. 12042
 /rpt_family="AT-rich"
 misc_feature
 /note="similar to Mus musculus EST W29262 (NID:q1309410)
 mc22c08.r1"
 repeat_region
 12057. 12081
 /rpt_family="AT-rich"
 repeat_region
 13116. 13146
 /rpt_family="AT-rich"
 repeat_region
 13371. 13454
 /rpt_family="L1"
 repeat_region
 13455. 13787
 /rpt_family="L1"
 repeat_region
 14714. 14943
 /rpt_family="MIR"
 repeat_region
 15410. 15462
 /rpt_family="AT-rich"
 repeat_region
 17142. 17190
 /rpt_family="(CA)n"
 repeat_region
 17798. 17989
 /rpt_family="MERL-type"
 repeat_region
 18156. 18391
 /rpt_family="L2"
 repeat_region
 20698. 20741
 /rpt_family="AT-rich"
 repeat_region
 21490. 21789
 /rpt_family="Alu"
 repeat_region
 21846. 21954
 /rpt_family="L2"
 repeat_region
 21957. 22046
 /rpt_family="GA-rich"
 repeat_region
 22125. 22147
 /rpt_family="AT-rich"
 repeat_region
 22566. 22843
 /rpt_family="Alu"
 repeat_region
 23662. 23943
 /rpt_family="Alu"
 repeat_region
 23946. 23988
 /rpt_family="(TAAAA)n"
 repeat_region
 24002. 24411
 /rpt_family="MERL-type"
 repeat_region
 24948. 24971
 /rpt_family="AT-rich"
 misc_feature
 25083. 25227
 /note="similar to Mus musculus EST W29262 (NID:q1309410)
 mc22c08.r1"
 repeat_region
 26227. 26432
 /rpt_family="MIR"
 repeat_region
 26469. 26755
 /rpt_family="MALR"
 repeat_region
 26756. 27051
 /rpt_family="Alu"
 repeat_region
 27052. 27155
 /rpt_family="MALR"
 repeat_region
 27157. 27613
 /rpt_family="MALR"
 repeat_region
 27897. 28473
 /rpt_family="MALR"


```

source      1. .387
            /organism="Homo sapiens"
BASE COUNT  143 a 51 c 91 g 102 t
ORIGIN

Query Match      10.6%: Score 356.8; DB 6; Length 387;
Best Local Similarity 96.8%: Pred. No. 3.4e-65;
Matches 364; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1882 gatttggcaactcaaccatgaacatgagcgcattcaatgaaataaaagctgcaaa 1941
      |||
DB 11 GATTTGGCACTCAACCTATGAACATGGCCCATTTCAATGGAATAAAGCTGCAAAA 70

QY 1942 aaaggaatcgcaagaacgctgttgcagaaacattggaagaagtacatgaagcccta 2001
      |||
DB 71 GAAGGAATCGCAGACGCTGTTGTCAGAACATTTGAGAGAGTACATGAGCCCTA 130

QY 2002 caaatgaatgacacattcgatgagatgctgtactcatttgaactttctatat 2061
      |||
DB 131 CAAATTAATGACACATTCGAAATGATGATGCTACTCATCTTGAAACTTTCTATAT 190

QY 2062 gaagagaagaataaagaatttcgacatgacatagaatgatagtatgaagtgatgat 2121
      |||
DB 191 GAAGAGAAGAATAAGACATTTGCACTCATAGACATGATGATGAGCGTGCTATGAT 250

QY 2122 gacttctgtatgctgatagaagaatgataaagaaccttgaactgataa 2181
      |||
DB 231 GAGTATTGTGATGCTGATGAAGAACATGATTTACTCAACCCCTTGAAACTGATGAA 310

QY 2182 acagataagattctcatgacttatttttgaacaataaattgttgaagaagctgct 2241
      |||
DB 311 ACAGATACATTCTCTATGACTTTATTTTGAACAAATATAACCTGATAAAGGCTGCT 370

QY 2242 gaaacccagaatagc 2257
      |||
DB 371 GAAACCCAGAAATATG 386

RESULT 12
AC007750 163681 bp DNA PRI 02-OCT-2000
LOCUS Homo sapiens BAC clone RP11-576116 from 2, complete sequence.
ACCESSION AC007750
VERSION AC007750.3 GI:6094634
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 163681)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 163681)
AUTHORS Cotton, M., Maupin, R., Hawkins, M. and Harkins, R.
TITLE The sequence of Homo sapiens BAC clone RP11-576116
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 163681)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 163681)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 163681)

```

```

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 22, 1999 this sequence version replaced g1:510389.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
Summary Statistics
Center project name: H_NH0576116
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPC1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.rgsen.com) or Pletier de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-178A14. Actual start of
this clone is at base position 1 of RP11-576116; actual end is at
base position 163681 of RP11-576116.
Location/Qualifiers
1. 163681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-576116"
/clone_lib="RPC1-11"
216..322
/rpt_family="MER81"
1271..1298
/rpt_family="(CAGCT)n"
1536..1694
/rpt_family="MIR"
2098..3518
/rpt_family="L2"
3614..3902
/rpt_family="A1u"
4090..4568
/rpt_family="MER1_type"
5128..5148
/rpt_family="AT_rich"
5263..5299
/rpt_family="(CA)n"
5417..5648
/rpt_family="L1"

```

repeat_region	/rpt_family-"L2"	7093. .7556	
repeat_region	/rpt_family-"MALR"	8702. .9187	
repeat_region	/rpt_family-"MIR"	9386. .9586	
repeat_region	/rpt_family-"L2"	10580. .10739	
repeat_region	/rpt_family-"Alu"	10792. .11064	
repeat_region	/rpt_family-"L1"	11136. .12912	
misc_feature	/note="match to EST A1077564 (NID:g4411972) G233q05.x1"	12894. .12899	
repeat_region	/rpt_family-"TTG)n"	12982. .13011	
repeat_region	/rpt_family-"TTA)n"	13109. .13133	
repeat_region	/rpt_family-"Alu"	13134. .13419	
repeat_region	/rpt_family-"Alu"	13833. .14130	
repeat_region	/rpt_family-"MIR1_type"	14437. .14573	
repeat_region	/rpt_family-"L1"	14742. .14867	
repeat_region	/rpt_family-"L1"	15098. .15234	
misc_feature	/note="match to EST A1922427 (NID:g5658391) W00b010.x1"	15444. .15460	
repeat_region	/rpt_family-"(T)n"	15449. .15475	
repeat_region	/rpt_family-"L1"	15476. .15695	
repeat_region	/rpt_family-"L1"	15713. .15993	
repeat_region	/rpt_family-"Alu"	15994. .16013	
repeat_region	/rpt_family-"(TAA)n"	16039. .16131	
repeat_region	/rpt_family-"L1"	16140. .16281	
repeat_region	/rpt_family-"L1"	16284. .16327	
repeat_region	/rpt_family-"(TG)n"	16328. .16394	
repeat_region	/rpt_family-"(CAT)n"	16396. .16576	
repeat_region	/rpt_family-"L1"	16625. .16997	
repeat_region	/rpt_family-"L1"	17020. .17274	
repeat_region	/rpt_family-"L1"	17270. .17603	
repeat_region	/rpt_family-"L1"	17644. .18442	
repeat_region	/rpt_family-"L1"	18444. .20187	
repeat_region	/rpt_family-"L1"	20260. .20562	
repeat_region	/rpt_family-"MER2_type"	20563. .20850	
repeat_region	/rpt_family-"Alu"	20851. .20910	
repeat_region	/rpt_family-"MER2_type"	21194. .21218	
misc_feature	/rpt_family-"(T)n"	21204. .21219	
misc_feature	/note="match to EST A1509439 (NID:g4598487) tr76a08.x1"	21204. .21219	
misc_feature	/note="match to EST A1582416 (NID:g4568313) tr97d11.x1"	21204. .21219	
misc_feature	/note="match to EST A1865901 (NID:g5530008) Wk88b06.x1"	21204. .21219	

	repeat_region	/note="match to EST A1583407 (NID:g4569304) tso9d01.x1"
		/rpt_family="MERL_type"
	repeat_region	/rpt_family="AT_rich"
	misc_feature	/note="match to EST A1682812 (NID:g4892994) wc66f03.x1"
	repeat_region	/rpt_family="Alu"
	repeat_region	/rpt_family="MALR"
	repeat_region	/rpt_family="MALR"
	repeat_region	/rpt_family="(CA)n"
	repeat_region	/rpt_family="MIR"
	misc_feature	/rpt_family="MERL_type"
	misc_feature	/note="match to EST T10929 (NID:g391083)"
	misc_feature	/note="match to EST A1682793 (NID:g4892975) wc66d05.x1"
	Query Match	9.0%; Score 303; DB 9; Length 163681;
	Best Local Similarity	100.0%; Pred. No. 1,2e-53;
	Matches 303; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	3063 ccagccttggggaacaaatgatgtgcacaagaqcttagatttgccctgtctcaataaag	3122
Db 149238	CCAGCGTTGGGGAACAATGATGTGTGCACAAAGCGTTAGATTGGCTTGTCATAAATAAG	149179
OY	3123 gaatttcgtagtgtttccaanaaatlccaacaaagaacaaalacaaaagtggta	3182
Db 149178	GAATTTTGTAGTGTGTTTCAAANAATTAATTCACAAAGAACAATACAAAAGCTGTACA	149119
OY	3183 attaccatcacattlccccaatcttgaactatcagaatgcgttattatgtatga	3242
Db 149118	ATTACTCATCACATTTGCCAAATCTTGACATTACGAATGCTCTTATTAGTAGAGAGA	149059
OY	3243 ttgaccttgatgaagatcttctaagaatcatcagttaaacatltaatatgat	3302
Db 149058	TTAGCAGCTTGATGAGATCTTTTAAATFACATTACAGCTTAACATTTAATATGATTAG	148999
OY	3303 attaatgtatcataatgctacagaagctgacataaagaalcaataaalgatgttact	3362
Db 148998	ATTAAATGATTATTATGCTACAGAACAGTACATTAAGAATCAATTAATGATTGTTACT	148939
OY	3365 ctg 3365	
Db 148938	CTG 148936	
RESULT	13	
LOCUS	150897	
DEFINITION	Sequence 7 from patent US 5643761.	
ACCESSION	150897	
VERSION	150897.1 GI:2472600	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 301)	
TITLE	Fisher,P.B. and Jiang,H.	
JOURNAL	Method for generating a subtracted cDNA library and uses of the	
FEATURES	Patent: US 5643761-A 7 01-JUL-1997;	
source	location/Qualifiers	
	1..301	
	/organism="unknown"	
BASE COUNT	110 a 61 c 56 g 74 t	
ORIGIN		

.
.
.

Result No.	Score	Query Match	Length	DB	ID	Description
1	711.6	21.1	1035	11	BF337466	BF337466 602035195
2	687	20.4	729	11	BF683236	BF683236 602305873
3	652.8	19.4	827	11	BF741116	BF741116 602631817
4	651.8	19.4	781	11	BF686405	BF686405 602143786
5	622.8	18.5	1041	11	BE882040	BE882040 601505326
6	618.4	18.4	1197	12	AK018602	AK018602 MUs muscu
7	531.4	15.8	536	10	A1718277	A1718277 as51606.x
8	520.4	15.5	537	10	A1806204	A1806204 w126c12.x
9	515.4	15.3	521	10	A1922705	A1922705 w011f08.x
10	511.2	15.2	552	10	AA134958	AA134958 z023f04.s
11	505.6	15.0	546	11	BE940626	BE940626 RC3-UT006
12	501.2	14.9	538	10	AM189584	AM189584 x108h12.x

C	13	480.4	14.3	484	10	AM152541	AM152541	xTf6bD2..x	AM152541	xTf6bD2..x
C	14	469.4	13.9	475	10	AM5895567	AM5895567	x012b12..x	AM5895567	x012b12..x
C	15	469.4	13.9	479	11	BF436606	BF436606	BF436606	BF436606	7F08d10..x
C	16	456.4	13.6	493	11	BC391418	BC391418	6024175007	BC391418	6024175007
C	17	456.4	13.6	499	10	AM779876	AM779876	hMh9c07..x	AM779876	hMh9c07..x
C	18	454.6	13.5	932	11	BC967630	BC967630	6028333060	BC967630	6028333060
C	19	445.6	13.2	481	10	AM580584	AM580584	CM4 - HT048	AM580584	CM4 - HT048
C	20	433.4	12.9	456	10	BE090628	BE090628	RC6 - BT072	BE090628	RC6 - BT072
C	21	432.2	12.8	446	10	BE090624	BE090624	RC6 - BT072	BE090624	RC6 - BT072
C	22	429.8	12.8	790	11	BF160649	BF160649	6017685449	BF160649	6017685449
C	23	429.4	12.8	452	10	BE090630	BE090630	RC6 - BT072	BE090630	RC6 - BT072
C	24	429.2	12.8	445	10	BE090627	BE090627	RC6 - BT072	BE090627	RC6 - BT072
C	25	405.6	12.1	585	10	AA123960	AA123960	mp78bD4..i	AA123960	mp78bD4..i
C	26	404.4	12.0	596	11	BF146945	BF146945	uy26c05..x	BF146945	uy26c05..x
C	27	369.6	11.0	447	10	AA099689	AA099689	ZK8bH09..S	AA099689	ZK8bH09..S
C	28	368.4	10.9	378	10	AA911194	AA911194	OK81d03..S	AA911194	OK81d03..S
C	29	346.8	10.3	425	10	AA135031	AA135031	z023f04..i	AA135031	z023f04..i
C	30	345.6	10.3	415	10	AA152349	AA152349	at78t108..x	AA152349	at78t108..x
C	31	342.8	10.2	482	10	AI190631	AI190631	RC-BT108	AI190631	RC-BT108
C	32	342	10.0	403	10	AI178346	AI178346	r135d07..x	AI178346	r135d07..x
C	33	335.6	10.0	349	11	H26598	H26598	y114a01..r1	H26598	y114a01..r1
C	34	330.4	9.8	520	10	BB284625	BB284625	BB284625	BB284625	BB284625
C	35	317.4	9.4	672	13	A0284992	A0284992	RPC111 - 75	A0284992	RPC111 - 75
C	36	310	9.2	776	11	BF163327	BF163327	6017718787	BF163327	6017718787
C	37	308	9.2	402	10	AA636698	AA636698	vp95c09..i	AA636698	vp95c09..i
C	38	299.6	8.9	477	11	C06405	C06405	C06405	C06405	C06405
C	39	285.4	8.5	699	10	BE306587	BE306587	601102907	BE306587	601102907
C	40	281.8	8.4	452	10	AM913323	AM913323	uF51f08..y	AM913323	uF51f08..y
C	41	281.8	8.4	526	10	AI152547	AI152547	u083bD6..i	AI152547	u083bD6..i
C	42	281.8	8.4	542	10	AM989970	AM989970	uF31e11..y	AM989970	uF31e11..y
C	43	269.2	8.0	377	11	Z20545	Z20545	H5AAACJNK	Z20545	H5AAACJNK
C	44	268.4	8.0	481	10	AA177334	AA177334	mt17910..i	AA177334	mt17910..i
C	45	253.4	7.5	671	11	BF124647	BF124647	6017610047	BF124647	6017610047

ALIGNMENTS

```

RESULT      1
LOCUS       BF337464
DEFINITION  BF337464 1035 bp mRNA EST 22-NOV-2000
ACCESSION   602035195F1 NC1:CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4183126
VERSION     5, mRNA sequence.
KEYWORDS    BF337464
SOURCE      BF337464.1 GI:11283715
            EST.
            human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Carniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1035)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapds@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: L14M9498 row: n column: 23
            High quality sequence start: 5
            High quality sequence stop: 695.
            Location/Qualifiers
                1..1035
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4183126"
                /clone_1b="NC1:CGAP_Brn64"
                /tissue_type="gliblastoma with EGFR amplification"
FEATURES
source

```


OY	1086	gcattccccgaagccagaactccagctcaggtcttcacaaatgaaagtltgccagccagcc	1125
Db	481	GCATCCCCGGAGACCAGAAACTCCAGCTCAGGCTTTACCAATATGGAATTGGCACCGAC-C	539
OY	1126	ttagaaggaaagaaalatacatcactctgctcccctacaagygagtggaaaaaccagatlgct	1185
Db	540	TTCGAAAGGGAACAATATCATCATCTG-CTCCTTACAGGAGGTGGAAAAACCAAGTGCGCT	598
OY	1186	gttttcatttgcgaagatcaccttaacaaaagaaagaaagctctgagcttgaaaagt	1245
Db	599	GTTTACATTGCCAAGAGATCACTTAAC-AAGAAGAAAAGCAATCTGAGCTTGAAAAATT	657
OY	1246	atagttcttgtlcaataaagtaagtaagctgtaagtgaacagctcttcgcgaagaagatltccaaca	1305
Db	658	ATAGTTCTTGTGCATAAGGTAAGTACTGTAAGTGAACAGCTCTTCGAAAGAGATTCACCA	717
OY	1306	tttttgaagaa	1316
Db	718	TTTTTGAGAA	728

SULT	3
BG741146	
LOCUS	BG741146
DEFINITION	827 bp mRNA
ACCESSION	602631817P1 NCI-CCAP-Sk03 Homo sapiens CDNA clone IMAGE:4776869 5'
VERSION	BG741146
KEYWORDS	mRNA sequence.
SOURCE	BG741146.1 GI:14051799
	EST.
	human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 827)
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	

Email: c9apds.r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM10630 row: b column: 06
High quality sequence stop: 763.

ATURES
source

Location/Qualifiers
1. .827

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:4776869"
/clone_1id="NCL_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NciI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5kb. Library constructed by Life
technologies. Note: this is a NCL_CGAP library."
BASE COUNT      254 a      186 c      204 g      183 t
ORIGIN

```

Query Match	19.4%	Score 652.8	DB 11	Length 827
Best Local Similarity	96.1%	Pred No. 8	2e-124	
Matches 798	Conservative	0	Mismatches 17	Indels 15
				Gaps 12

Oy 426 cggcggccctctggccgccgacatgaacccctgaactcggacttgcctctccatc 485
 1 cggcggccctctggccgcc -ccgctacatgaacccctgaactcggacttgcctctccatc 59
 Oy 486 gtttagaacgtctatgttgatatctcccaactctgaactcctcttaag -cccaactctg 544
 |||||

Db	60	GTTTGACAAAGCGTCATGATGATGATATCTCCAACTGCTGAACCTCCTTCAGTCCCACTGG	119
OY	545	tgacaaagcttccagtttagtagagacgctcttgataaagtgcagagagagaaactglttgacaa	604
Db	120	TGSCAAGCTTTAGTTAAGAGACGTTCTTGATTAAGCTGATGAGAGAACGTTTGACAA	179
OY	605	ttgaagacaaaaccggatltgctgctgcagaaaacaatlguaatgaaatcaggtgtaagag	664
Db	180	TTGAAAGCAAAACCGGATTTGCTGTGCGAAGAAACAAATGAAATGAATCACTGTGTAAGAG	239
OY	665	agctaacaaaagatlttgacagaaaagaaactggtctctgcattcttcgaatgttc	724
Db	240	AGCTACTAAAAAGGATTTGTGCACAAAGAAAACCTGGTTCTCTGCATTTCTGAATTTTTC	299
OY	725	gtcaacaacagaaacaatlgaaactgtgtccaaagatltacaagcgtcttgatg--ctcaagaag	782
Db	300	GTCAAACAGGAACAAATGAACTTGTCCACAGACTTAACAGCGCTTGATTGGCTCAGAAAG	359
OY	783	caatgcagagattgagaattatcacaaagttgaaatgtcccaagtgtgaagaagcaactct	842
Db	360	CAATGCGAGATTTGACAAATTTATCACAACTGATGTGTCCTCAAGTGCAGAGCAACTTCT	419
OY	843	ttcaacacaggttcagccaactcttga--gaagaggtctggagcatgagaataaacat	901
Db	420	TTCAACCAACGTTTACGCCAATCTGGACGCAAGAGGCTCTGGGGCATGGAGATTAACCAT	479
OY	902	cagaatcatcttttga--gattctcttgagtttcaagatcaagacaagaatttga--g	958
Db	480	CAGAAATCATCTTTTGGACAGATTCTTCTGATTTTCAGAAATCGAACACAAAGTTTGGCAGCG	539
OY	959	aaggaagatgtcagctgctttagatgaaagtcttga--cataacagcaacatgggcagtat	1017
Db	540	AAGGAAGTGTCACTGCTGTAGATGAAGAAAGCTTGGACCAATACAGAACATGTGGCAGATGAT	599
OY	1018	tcaagacacatgggaagttaattacagatga--gagaatltggcagcaaaagatccscgga	1076
Db	600	TCAAGCAACCATGGGAAGTGTATTGATGAGACGAGATGTTGGAGCAAAAGCATCCCCGGA	659
OY	1077	gcacgaactccaagctcagcgtcttaccaaatg--aaagtgtccacgcacgctcttgaagag	1134
Db	660	GCCAGAACTCCAGCTCAGCGCTTACCAAAATGGCACAGTTGGCCAGGCCAG--CTTGGCAGGG	718
OY	1135	aagaataatcatcttgctccctacagggagttgaaaaaccagatgtgctgtttacat	1194
Db	719	AAGACTATCATCTATCTTG--CTCCCTACAGGAGATGCAAAAAACAGAAAGTGGCTGATTACAT	777
OY	1195	-ggcaagatcatcttagacagagagaaaaaacatctgaagcttgaagaag	1243
Db	778	AGCCACGGATCTCTTAAACAAGAAACAAACATCTTGAAGCTCGAAGAG	827

RESULT	4
BF686405	
LOCUS	781 bp mRNA
DEFINITION	6021343786f1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304805.5'
ACCESSION	BF686405
VERSION	BF686405.1 GI:1971813
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (pages 1 to 781)
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE	NIH-MGC http://mgc.nhl.nih.gov/ , National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strassberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.

QY 3253 atgaagatctttaaatactatcaqtaaacattatgatgatgaatgcat 3312
|||||
Db 116 ATGGAAGATCTCTTTAAATACTATCACTTAAACATTTAATATGATTAATGATAT 57
QY 3313 tcaatagctacgaagactgacataagaatacaataatgatgtttactctg 3365
|||||
Db 56 TCATATATGCTACAGAACTGACATAGAAATCAATAATGATGTTTACTCTG 4

RESULT 8
A1806204/c 537 bp mRNA EST 19-DEC-1999
LOCUS wF26c12.x1 Soares.NFL.T.GBC.S1 Homo sapiens cDNA clone

DEFINITION IMAGE:2356726 3', mRNA sequence.
ACCESSION A1806204
VERSION A1806204.1 GI:5392770

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 537)
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Insert length: 729 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 444.

FEATURES

source

Location/Qualifiers

1..537

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2356726"

/clone_lib="Soares_NFL.T_GBC.S1"

/lab_host="DH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not 1; Site: 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung MDHL19W, testis NRT, and B-cell
NCL-GAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.

BASE COUNT 157 a 93 c 84 g 203 t
ORIGIN

Query Match 15.5%; Score 520.4; DB 10; Length 537;
Best Local Similarity 99.6%; Pred. No. 1.2e-96;
Matches 532; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2832 caaggaatatttgcgaagcattacaagaataccaccataacttccttgcga 2891
|||||
Db 537 CAAGGAATATTGGCCAGCACTTACAGAAATTAACCCAT-ACATAATACCTTCCTTGCAA 479
QY 2892 aaacgcaagtgtcagcctgtctcgggaagataatccatgtaattggaagaatgcat 2951
|||||
Db 478 AAACGCAAGTGTCTAGCCTGTCTGGGGAAGATATCCATGTAATTGGAGAAATATCATCA 419
QY 2952 cgtcaataagaccagaattcaaggaattacattgtaagagaacaagaacatgca 3011
|||||
Db 418 CGTCAATATGAGCCGAGAAATTCAGGAACCTTACATTTGTAAGAGAAACAAACACTGCA 359
QY 3012 aaagaagtgtcgcgaactcaataaattgtgaaatcatctgcgaattgtgcgaagcttg 3071

|||||
Db 358 AAAAGACTGTCCGACATCATCAATTAATGCTGAATCATTCGCAAGTGCAGGCTTG 299

QY 3072 gggacaatgatggtgcacaaagccttagattgccttgcctcacaataagaaatttgt 3131
|||||

Db 298 GGGAAACATATGATGCTGACAAAGCCTTAGATTGCCCTTGCTCAAAATAAGAAATTTGT 239

QY 3132 agtggctttaaataatattcaacaagaatacaataaagtggtaattactat 3191
|||||

Db 238 AGTGCTTTCAAAATATTCAACAAAGAAACATTAACAAAGCGGTACAAATTACCTAT 179

QY 3192 cacaatcccaatctgacattcagaatgctgttattagatgataagattgacatt 3251
|||||

Db 178 CACATTTCCCAATCTTACATATTAGAAATGCTGTTATTAGATGATGAGATTAACACTT 119

QY 3252 gatgaagattctttaaatactatcaglttaaacattatgatattgatgta 3311
|||||

Db 118 GATTGAAGATCTCTTTAAATACTATCATGTTAAACATTATGATTAATGATTAATGTA 59

QY 3312 ttcattatgctacgaagactgacataagaatacaataaagtattgttactctg 3365
|||||

Db 58 TTCATTATGCTACAGAACTGACATAGAAATCAATAATGATGTTTACTCTG 5

RESULT 9
A1922705/c 521 bp mRNA EST 01-SEP-1999
LOCUS w011f08.x1 NCI-CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2455047 3',
mRNA sequence.

DEFINITION A1922705
ACCESSION A1922705
VERSION A1922705.1 GI:5658669
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 521)
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrr/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 410.

FEATURES

source

Location/Qualifiers

1..521

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2455047"

/clone_lib="NCI-CGAP_Pan1"

/tissue_type="adenocarcinoma"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: PCMV-SPORE; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 154 a 90 c 82 g 195 t
ORIGIN

Query Match 15.3%; Score 515.4; DB 10; Length 521;
Best Local Similarity 99.8%; Pred. No. 1.3e-95;
Matches 516; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2849 agccttaagaataaacaccataacttaacttccttgcgaagaatgcaatgtag 2908
|||||
Db 521 AGCATTAACAAGAAATTAACCATCACTAATAACTTTCCTTGCAAAACGACGTGTCTAG 462

OY 2909 cctgtcttgggaagatataccatgttaattgagaagaatgcaatcactcaatatacaccacag 2968
 |||||||
 Db 461 CCTGTTTGGGGAAGATATCCATGTAATTGAGAAATGCATCAGCTCATATGACCCAG 402
 OY 2969 aattcaagaacttaccatttgaagaagaacaagaacagcagcagaagaagtgcagact 3028
 |||||||
 Db 401 AATTCAAGAACCTTACATTGTAAACGAAACAAACAAACCTCCAAAGAGAGTCCCGACT 342
 OY 3029 atcaaatatggtgaatcatctcgaatctgagcgaatctggggaacaatgatgtgc 3088
 |||||||
 Db 341 ATCAATATTAATGTAATTCATCTGCAATATGCGCAGGCTTGGGGAACATGATGTGC 282
 OY 3089 acaaaagcttagatttgccttgcctcaaaalaaggaatttctagtgttccaataa 3148
 |||||||
 Db 281 ACAAGGCTTGATTTGCTTGTCTGCAAAATAGGAATTTGTAGTGTTCATCAAAATA 222
 OY 3149 attcaagaagaagaataacaaagaagtggtgagaattactatcacatttcccaacttg 3208
 |||||||
 Db 221 ATTCAACAAACAAACAAACAAACAAAGTGGTGAATTCATTCATTCATTCCTTG 162
 OY 3209 actatcgaatgctgttatttattgagatgaggaattacacattgaaagattcttta 3268
 |||||||
 Db 161 ACTATTCGAATGCTGTTTATTATTAGTATGAGGATTCAGCACTTGATTCAGATTTCTTTA 102
 OY 3269 aaatactatcgaatcaaatcaatataatataatgataatgattatcattatctacagaa 3328
 |||||||
 Db 101 AAATACATCATGATTAACATTTAATATGATTATGATTAACTATTCATTCATTCATTCAGAA 42
 OY 3329 ctgacataagaatcaataaataatgatttcttactctg 3365
 |||||||
 Db 41 CTGACATTAAGATCAATTAATGATTCTTACTCTG 5

RESULT 10
 AAI34958/c 552 bp mRNA EST 30-NOV-1996
 LOCUS AAI34958
 DEFINITION zc3104.s1 Strata gene colon (#937204) Homo sapiens cDNA clone
 IMAGE:587743 3', mRNA sequence.
 AAI34958
 AAI34958.1 GI:1696041
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 552)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chippelli, B.,
 Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
 M., Hultman, M., Kucada, T., Lacy, M., Le, N., Mardis, E., Moore,
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevas, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Wrevas, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 TITLE JOURNAL MEDLINE
 COMMENT 97044478
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 278.
 Location/Qualifiers
 FEATURES
 Source
 1..552
 /organism="Homo sapiens"
 /db_xref="GDB:4620045"
 /db_xref="taxon:9606"
 /clone="IMAGE:587743"
 /clone_id="Strata gene colon (#937204)"

/issue_type="tumor"
 /cell_line="T84 carcinoma cell line"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: colon. Vector: pBluescript SK-; Site: 1;
 EORI: Site: 2; XhoI: Cloned unidirectional. Primer:
 01190 dt. T-84 colonic epithelial cell line. Average
 insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor
 sequence: 5' GAATTCGGCAGCG 3' -3' adaptor sequence: 5'
 CTCGACGTTTCTTTTCTTTT 3'

BASE COUNT 157 a 97 c 90 g 206 t 2 others
 ORIGIN
 Query Match 15.2%; Score 511.2; DB 10; Length 552;
 Best Local Similarity 98.0%; Pred. No. 9.3e-95;
 Matches 527; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

OY 2828 aaacagaagaatattgcaagaatcacaagaataacacacataaacttccctt 2887
 |||||
 Db 551 AAACCCAGGGAATATTGCCAGGCAAT-CCAGGATACCCCATCAATTAACCTTCTTT 493
 OY 2888 gcaaaactgcaagtgctgactgctgcttctgggaagaatatacattgaatgaagaatgc 2947
 |||||||
 Db 492 GCAAAACCTGCAAGTGTCTAGCCTGTTCTGGGAAGATATCATGTAATTGCAAAATGC 433
 OY 2948 atcacgcaatagaccacgaatcaagaacttatacattgtaagaagaacaaagac 3007
 |||||||
 Db 432 ATCAGCTCATATGAGCCCGCAATTCAGGACTTTAATTGTAAGCAAAACAAACAC 373
 OY 3008 tgcagaagaagtgctgcaatcaaaalaagaatgcaatcagcaatgagcag 3067
 |||||||
 Db 372 TGCAAAAGAGTGTGCCGACTATCAATTAATGCTGAATCAATCATCTGCAATGTGGCAGG 313
 OY 3068 ctggaggaacaatgagtgacacaaagcttagatttgccttgcctcaaaalaagaatt 3127
 |||||||
 Db 312 CTGGGGAACAAATGATGTGACCAAGGCTTACATTTGCTTCTCAAAATAGCAATT 253
 OY 3128 ttgtagtgcttcaaaaataatcaacaagaagaacatacaaaagtggtgagaaatc 3187
 |||||||
 Db 252 TTGTAGTGTGTTTCAAAAATTAATTCACAAAGAACAAATGAGGCTGAATTAC 193
 OY 3188 ctatcaatctcccaactctgactatctcgaatgctgcttatttgaatgaagaatgc 3247
 |||||||
 Db 192 CTATGACATTTCCCAATCTTGACATTCAGAAATGCTGTTTATGAGAGGATTAGC 133
 OY 3248 actgattgaagaattctttaaataactacatcagttaaacatttaataatgattgaa 3307
 |||||||
 Db 132 ACTTGATGGAAGATCTTTTAAATTAATCACTTAACATTTAATATGATTATGATTAA 73
 OY 3308 tgaattcattatgctacagaactgacataagaatcaataaataatgattgattactctg 3365
 |||||||
 Db 72 TGATTCATTATGCTACAGAACTAGATTAAGTAATTAATGATTGTTTACTCTG 15
 RESULT 11
 BE940626 546 bp mRNA EST 02-OCT-2000
 LOCUS BE940626
 DEFINITION RC3-UT0064-170800-021-e06 UT0064 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE940626
 VERSION BE940626.1 GI:10470130
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 546)
 Nagai, M.A., Garcia, C., Verjovski-Almeida, S., Briones, M.R.,
 Dias Neto, E., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bal, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
MEDLINE
COMMENT

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC3-UT0064-170
800-021-e06&t3=2000-08-17&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 53
High quality sequence stop: 542.
Location/Qualifiers

FEATURES

source
1..546
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0064"
/dev_stage="Adult"
/note="Organ: uterus; tumor; Vector: puc18; Site: 1; Smal;
Site: 2; Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 205 a 86 c 102 g 153 t
ORIGIN

Query Match 15.0%; Score 505.6; DB 11; Length 546;
Best Local Similarity 98.9%; Pred. No. 1.3e-93;
Matches 530; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 2778 ttaataagatttggaaattacagatcgaagtataatgaa-aagaaatgaaccaga 2836
DB 11 TCATCAGATTTGGATTACAGATTAAGTAAATGACAAAGAAAGAAACCAACCA 70
QY 2837 gaataatccagattcaagaataaccataactccttcgcaaaact 2896
DB 71 GAAATTA-TGCCAAGCATTAACAGATTAACCATCACTATTACTTCTTGCACAAACT 129
QY 2897 gcaagtgtcgtcgttctcgtggaagatataccatgtaattgagaataatgacacgtca 2956
DB 130 GCAGTGTCTACCTGCTTCTCGGGAAGATATCCATGTAATGAGAAATGATCAGCTCA 189
QY 2957 atataagcccaagaattcaagaactttacattgttaagagaagaagaagacgtcgaaga 3016
DB 190 ATATGACCCCAAAATTCAGGAACCTTTACATTGTAAAGAAACAAAGCACATGCAAAA 249
QY 3017 agtgcacacatcaataaataatgtaaatcatctgcaaatgtgcaggtctggaaga 3076
DB 250 ACTGTGCGACATTCATAATATGTTGTAATCATCTGCAAAATGTGGCCAGGCTTGGGGA 309
QY 3077 caatgagtgacaaaggcttagattgtcctgtctcacaataaagaaatttgaatg 3136
DB 310 CAATGATGTGCAAAAGCTTAGATTGCTGTCTCAAAATTAAGGAATTTTGTAGTGG 369
QY 3137 ttttcaaaataattcaaaaagaataatacaaaaagagggtgaattactatcaact 3196
DB 370 TTTTCAAAATATTCACAAAGAAACAAATTAAGGAATTTTGTAGTGG 429
QY 3197 ttcccaatctgaactcaagaatgctgttattagatgaagattgacactgattg 3256
DB 430 TTCCCAATTTGACTATTCAGATGCTGTTTATTAGTATAGGATTGACATTTGATGATG 489
QY 3257 aagattcctttaaataactatcagttaaacattataatgattatgattatgatt 3312
|||||

Db 490 AAGATCTTTAAATACACTACAGTTAAACATTAAATGATTGATTACTGAT 545

RESULT 12
LOCUS AM189584/c 538 bp mRNA EST 22-NOV-1999
DEFINITION x108h12.x1 NCI_CGAP_U4 Homo sapiens cDNA clone IMAGE:2675687 3'
similar to TR:095786 095786 RNA HELICASE. ; mRNA sequence.
ACCESSION AM189584
VERSION AM189584.1 GI:6464064
KEYWORDS EST.
SOURCE human.
ORGANISM human.

REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Tissue Procurement: Christopher Mokaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.linl.gov/bdrp/image/image.html
Seq primer: -400p from Gibco
High quality sequence stop: 175.
Location/Qualifiers

FEATURES

source
1..538
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2675687"
/clone_lib="NCI_CGAP_U4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site: 1; SalI;
Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"
BASE COUNT 157 a 92 c 94 g 195 t
ORIGIN

Query Match 14.9%; Score 501.2; DB 10; Length 538;
Best Local Similarity 97.4%; Pred. No. 1e-92;
Matches 520; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 2832 caagagaataatgtcgaacattcaagaataaccataactccttcgtgcaa 2891
DB 538 CAAGGAATATTTGCGACGACATTTACCAACATTAACCCATCACTAAT-ACCTTCTTGCCAA 480
QY 2892 aaactgcagtgctgagcgttctcgtggaagatatccatgtaattgagaataatgcatca 2951
DB 479 AAACGTGAGTGTGCTAGCTTTCTGCGGAAGATATCCATGTAATTGAGAAATGACATCA 420
QY 2952 cgtcaatgccccaggaattcaaggaactttacattgtaagagaagaagaagacgtcga 3011
DB 419 COTCAATATGACCCGACGATTCAGAAATTCATTGTAAGAGAAACCAAGACACTGCA 360
QY 3012 aagaagatgtccgactatcaataaataatgtaaatcatctgcaaatgtgcaggtctg 3071
DB 359 AAAGAGTGTGCGACGACATTCATAATTAATGTTGAAATCATCTGCAAAATGTGCGAGCTTG 300
QY 3072 gggacaatgtagtgcgaagaaggttagattgtcctgtctcacaataaagaattttgt 3131
DB 299 GGGACCAATGATGTGCAAAAGCTTAGATTGCTGCTGTCACAAATTAAGGAATTTGT 240
|||||
